VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Edited by Richard Schmalensee

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

From the Editor

Richard Schmalensee

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

From the Editor

Firms with market power engage in a variety of business practices that harm their rivals. Under what circumstances should the antitrust laws condemn these practices because they will harm consumers? This long-standing question is being discussed with renewed intensity both in the European Community and in the United States. The European Commission's Directorate-General for Competition (DG COMP) has been working on a document explaining its views on this question for more than a year. In December 2005, it released the draft of a discussion paper (Discussion Paper) on which it has sought comments.[1] Meanwhile, in the United States, the Antitrust Modernization Commission and, more recently, the U.S. Federal Trade Commission (FTC), have focused on this question. In both jurisdictions, the debate has been stimulated in part by controversial court decisions concerning so-called loyalty rebates.[2]

Our first issue of 2006 begins with a symposium that contributes to this discussion. Alden Abbott and Michael Salinger, both with the FTC, begin by examining tying-a practice that is often treated as a restraint of trade under the U.S. antitrust laws and as an abuse of dominance under Article 82 of the EC Treaty. They question the approach taken both by the courts and by their fellow economists. Professor Herbert Hovenkamp, of the University of Iowa, College of Law, looks at predatory pricing.[3] He observes that the U.S. courts tend to leave practices they do not understand to juries. He argues that as the courts become more familiar with the new variants of predatory pricing claims, such as those involving loyalty rebates, they will limit the ability of plaintiffs to put these claims in the hands of juries, much as they did for traditional predatory pricing allegations.

1 European Commission, DG Competition Discussion Paper on the Application of Article 82 of the Treaty to Exclusionary Abuses (Dec. 2005), *available at* http://europa.eu.int/comm/competition/antitrust/others/discpaper2005.pdf.

2 *See* Symposium, *A Symposium on Loyalty Rebates*, 1(2) Competition Pol'y Int'l 89 (2005).

3 *See* H. Hovenkamp, The Antitrust Enterprise: Principle and Execution (2006).

Richard Schmalensee is Professor of Economics and Management at MIT and John C Head III Dean of the Massachusetts Institute of Technology's Sloan School of Management.

The symposium then switches to Europe. Two articles by leading practitioners of EC competition law consider the Commission's recent Discussion Paper. Bill Allan, of Linklaters, provides an overview and critique of the Commission's Discussion Paper. He argues that the Discussion Paper does not go far enough. In his view, the analysis of exclusionary abuses should examine whether they significantly lessen competition in a way that can be remedied by the application of the antitrust laws. Frank Montag and Alicia Van Cauwelaert, of Freshfields Bruckhaus Deringer, focus on the Commission's examination of refusals by dominant firms to supply other firms with access to their physical or intellectual property. They argue that the Discussion Paper improperly widens the scope of refusal to supply abuses and advocate a "no economic sense" test to limit the application of this abuse.

Lastly, we present a paper by a group of economists who were asked by DG COMP's Office of the Chief Economist to provide their advice on how to apply competition law. The Economic Advisory Group for competition policy (EAGCP), consisting of Professors Jordi Gual, Martin Hellwig, Anne Perrot, Michele Polo, Patrick Rey, Klaus Schmidt, and Rune Stenbacka, argue for an effects-based analysis founded on economic theory and empirical evidence.

The next part of this issue consists of an article on the "trading services" industry, which consists of financial exchanges and other businesses that facilitate the trading of financial instruments. Bernhard Friess and Sean Greenaway of DG COMP examine this industry and find that, given the inherited structures, both regulation and monitoring by competition authorities is warranted to deal with a variety of market failures.

Our Autumn 2005 issue had a provocative exchange on vertical restrictions between FTC and U.S. Department of Justice economists Cooper, Froeb, O'Brien and Vita in one camp and Professors Scherer and Winter in another camp.[4] The discussion continues in this issue.

In this third issue of *CPI*, we start a new feature-short articles that examine recent significant legal decisions around the world. William Rooney, of Willkie Farr & Gallagher, looks at the U.S. Supreme Court's decision in *Volvo Trucks*,[5] a case that concerned price discrimination under the U.S. Robinson-Patman Act. Shaun Goodman, of Cleary Gottlieb Steen & Hamilton, reviews the European Court of First Instance's decision in *General Electric/Honeywell*.[6] That case concerned an appeal of the European Commission's decision to block the merger.

4 *See* Colloquy, *A Colloquy on Vertical Restrictions,* 1(2) COMPETITION POL'Y INT'L 45 (2005).

5 Volvo Trucks N. Am., Inc. v. Reeder-Simco GMC, Inc., 546 U.S. ___, 126 S.Ct. 860 (2006).

6 Commission Decision 2004/134/EC, General Electric/Honeywell, 2004 O.J. (L 48) 1.

While there is still, as these articles show, room for respectful disagreement over the application of the antitrust laws to particular cases, there has emerged widespread agreement that enhancing long-run consumer welfare is the singular goal of competition policy. The debate, therefore, can focus on how to achieve that objective. The emergence of this agreement was hastened, at least, by the highly influential article by Robert Bork, which we reprint in this issue, along with an introduction by Chief Judge Douglas Ginsburg of the U.S. Court of Appeals for the DC Circuit. Judge Ginsburg observes that scholars have seriously questioned Bork's position that promoting consumer welfare was the intent of the U.S. Congress in adopting the Sherman Antitrust Act of 1890. Nevertheless, Bork's view that consumer welfare should be the sole objective of the antitrust laws has been widely embraced by the courts and others and has led to efficiency-enhancing decisions.

On behalf of the journal's readers and its editorial team, I am delighted to extend my thanks to all the contributors of this issue.

Richard Schmalensee
Editor-In-Chief

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Learning from the Past: The Lessons of Vietnam, IBM, and Tying

Alden F. Abbott and Michael A. Salinger

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Learning from the Past: The Lessons of Vietnam, IBM, and Tying

Alden F. Abbott and Michael A. Salinger

With a major set of hearings scheduled in the United States on the antitrust treatment of single-firm conduct, economists have an opportunity to provide analysis that informs policy. Yet, the opportunity will be lost if economic analysis does not provide insights into how to distinguish anti-competitive from pro-competitive behavior. We argue that the economics literature on one type of single-firm conduct—tying—has been less influential than it should have been, and examine whether there are lessons to be learned from that failure. We argue that the two principal causes are 1) the almost complete neglect of competitive tying (while focusing heavily on anticompetitive tying) and 2) an excessive reliance on theory alone.

Alden F. Abbott is Associate Director for Policy and Coordination in the Bureau of Competition at the Federal Trade Commission. Michael A. Salinger is Director of the Bureau of Economics at the Federal Trade Commission and on leave as Professor of Economics at the Boston University School of Management.
The authors thank Liz Callison, Mark Frankena, and Josh Soven for helpful comments. The views expressed here are those of the authors and do not necessarily reflect those of the U.S. Federal Trade Commission or any individual Commissioner.

I. Introduction

During the administration of Lyndon Johnson, the United States undertook two major conflicts that lasted much longer than anticipated, ended in withdrawal in the face of defeat, and left it reluctant for decades to exercise its powers. One conflict was Vietnam; the other was the monopolization suit against IBM. Just as, several decades later, the United States did overcome the "Vietnam Syndrome" and sent its troops abroad, the U.S. antitrust authorities have resumed challenges of dominant-firm conduct. Within the last decade, the U.S. Department of Justice (DOJ) has brought such suits against Microsoft,[1] American Airlines,[2] and the Visa and Mastercard networks,[3] to name a few.

Continuing with the Vietnam analogy, virtually everyone agrees on the importance of learning the "lessons of Vietnam." As is evidenced by debates over whether the choice and conduct of current conflicts demonstrate learning from, or repetition of, past mistakes, exactly what those lessons are is less clear. Similarly, the antitrust community went through an attempt to learn what might be termed "the lessons of IBM," recognizing that IBM, flawed as the case was, is synecdoche for a problematic history of monopolization, monopoly leveraging, and monopoly maintenance cases.[4] With the revival of cases about the behavioral limits on a firm with a dominant market share, agency-sponsored analyses are, or soon will be, occurring on both sides of the Atlantic. In the United States, Federal Trade Commission (FTC) Chairman Deborah Platt Majoras and DOJ Assistant Attorney General Thomas Barnett have announced a major set of hearings on Section 2 of the Sherman Act.[5] The European Commission has recently released a report on Article 82 that broaches many of the same issues.[6] In these assessments, a good overarching question to ask is whether we have learned the lessons of IBM.

A proposition that is almost as widely held as the necessity of learning the lessons of Vietnam, is that antitrust enforcement must be informed by solid economic analysis. The examinations of policy toward unilateral conduct provide economists with an opportunity to provide useful input. Law enforcers without

1 United States v. Microsoft Corporation, 253 F.3d 34 (D.C. Cir. 2001) [hereinafter *Microsoft*].

2 United States v. AMR Corporation, 335 F.3d 1109 (10th Cir. 2003).

3 United States v. Visa U.S.A., Inc., 344 F.3d 229 (2nd Cir. 2003).

4 Franklin M. Fisher et al., Folded, Spindled and Mutilated: Economic Analysis and U.S. vs. IBM (1983).

5 Press Release, U.S. Federal Tqrade Commission, FTC and DOJ to Host Joint Public Hearings On Single Firm Conduct as Related to Competition (Nov. 28, 2005), *available at* http://www.ftc.gov/opa/2005/11/unilateral.htm.

6 European Commission, DG Competition discussion paper on the application of article 82 of the treaty to exclusionary abuses (Dec. 2005), *available at* http://europa.eu.int/comm/competition/antitrust/others/discpaper2005.pdf.

formal training in economics recognize the inherent difficulty of distinguishing abusive behavior from either aggressive competition or legitimate strategies to reap the rewards of legally obtained market power, and appear eager for economists to provide insights that lead to practical legal rules. There is no guarantee, however, that the economics profession will capitalize on this opportunity. Indeed, we believe that before giving the advice that will be solicited on single-firm conduct, economists should do their own reckoning of the past. Have antitrust authorities and courts made bad decisions because they ignored the clear, sound advice from the economics profession, much as Lyndon Johnson ignored advice that escalating the war without raising taxes at a time of full employment would likely lead to inflation? Or have economists simply failed to provide law enforcers with the analysis they need to make good decisions?

Law enforcers want clear, simple rules. Judge (and Professor) Frank Easterbrook has advocated the use of basic filters to evaluate particular forms of conduct under the antitrust laws, and Professor Richard Epstein has proposed the use of simple rules in law enforcement.[7] Building on these insights, case filters in the form of "simple rules" might be particularly attractive for use in the evaluation of unilateral conduct. Simple rules would tend to reduce the degree of antitrust-specific business uncertainty that deters efficiency-enhancing unilateral behavior, and thereby promote social welfare.[8] Moreover, simple rules would cabin judicial discretion and thereby reduce the costs and uncertainty associated with judicial evaluation of unilateral business conduct.

Arguably, the model of a simple rule in the Section 2 context is the U.S. Supreme Court's *Brooke Group* holding[9] that required a showing of both (1) below cost pricing and (2) the likelihood of recoupment to support a finding of single-product price predation. Notably, the *Brooke Group* rule has eliminated a great deal of costly litigation and has reduced business uncertainty in one area of conduct. *Brooke Group* does not eliminate all false acquittals—indeed, various theoretical economic models demonstrate how, given certain assumptions, above-cost single-product price cuts can be anticompetitive.[10] Implicit in the

7 *See* Frank Easterbrook, *The Limits of Antitrust*, 63 Tex. L. Rev. 1 (1984) (*reprinted in* 1 Competition Pol'y Int'l 179-215 (2005)). More generally, Professor Richard Epstein advocated the use of simple legal principles to improve the treatment of complex problems in modern society (problems that typically have been dealt with by very complex legal regimes). *See* Richard Epstein, Simple Rules for a Complex World (1997).

8 Although collaborative conduct involving competitors also might benefit from the application of simple rules, such norms are even more important in the context of single firm conduct. Restrictive joint behavior is more likely to have pernicious effects on competitive rivalry than unilateral behavior, undertaken by a firm that is seeking to outdo its rivals.

9 Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993) [hereinafter *Brooke Group*].

10 For a general discussion of above cost predatory pricing scenarios, *see, e.g.*, A. Edlin, *Stopping Above-Cost Predatory Pricing*, 111 Yale L.J. 941 (2002).

Brooke Group standard, however, is the assumption that any harm stemming from a failure to prosecute the rare legitimate predation case is more than outweighed by the benefits of avoiding unsound enforcement actions and costly business uncertainty that would occur in the absence of this simple rule.[11] Simple rules can be desirable even if they do not yield the correct result in every case in which they are applied—they merely need to satisfy the criterion that overall welfare will be higher in the presence of the rule than in its absence.[12]

The challenge for economists is to help formulate rules that are, to quote Einstein, "as simple as possible, but no simpler." Doing so might be relatively easy for some types of cases. An example might involve a firm's manipulation of government processes to delay or deter entry (as in the FTC's *Orange Book* cases[13]) or otherwise obtain market power (as in *Unocal*[14]). Another example might be a firm's commission of intentional torts that appear to allow it to obtain or maintain market power. In both those categories, the conduct would appear to be "always or almost always" inefficient, and thus, unlikely to yield false positives and deter welfare-enhancing conduct. The "cheap exclusion" framework (referring to behavior that lacks any plausible efficiencies and is relatively inexpensive to undertake) developed by former FTC official Susan Creighton and colleagues[15] is a methodology that subsumes these examples. It suggests a way of developing simple rules to condemn behavior based on the behavior's lack of efficiency justifications and relatively low costs to the alleged predator.

Of course, it is the harder classes of practices, such as tying and exclusive dealing, where law enforcers are most in need of help from economics. The practices are widely used and are usually efficiency-enhancing, but in some situations, their primary purpose may be to contribute to monopoly maintenance rather

11 Per se prohibitions, such as the per se rules against naked horizontal price fixing, also implicitly assume that the error costs stemming from those rules (in those cases, the harm from rare false positives) are more than outweighed by the rules' benefits (deterrence of harmful behavior and ease of administration).

12 Using the somewhat more abstract language of decision theory, the objective in formulating simple rules is to minimize the sum of expected error costs and enforcement costs.

13 For a more detailed description of these cases, *see* Bureau of Competition, U.S. Federal Trade Commission, Overview of FTC Antitrust Actions in Pharmaceutical Services and Products 3-4 (Jun. 2005), *available at* http://www.ftc.gov/bc/050802antitrustpharmprods.pdf.

14 For a more description of this matter, *see* Statement of the Federal Trade Commission in the Matter of Union Oil Company of California, Docket No. 9305, and Chevron/Unocal, File No. 051-0125 (Aug. 2005), *available at* http://www.ftc.gov/opa/2005/06/chevronunocal.htm.

15 *See* Remarks by Susan Creighton (at the time, FTC Bureau of Competition Director), *Ranking Exclusionary Conduct,* ABA Section of Antitrust Law Fall Forum, Nov. 15, 2005, *available at* http://www.ftc.gov/speeches/creighton/051115conduct.pdf.

than to generate efficiencies, as the courts concluded in *Microsoft*[16] *and* Dentsply.[17] The challenge is how best to distinguish, albeit imperfectly, anticompetitive instances of these practices from those that are pro-competitive or at least competitively neutral.[18] There is a presumption that these practices are also per se legal for firms without market power. While, for firms with market power, per se legality would be simple, but it is too simple.

In these situations where law enforcers need something more nuanced than per se rules to ascertain when behavior is likely to be anticompetitive, economists should acknowledge that the existing economics literature falls short of giving law enforcers what they need. An important source of the problem is that the modern economics literature on single-firm conduct starts from the assumption that a firm has market power and then analyzes if and when certain behavior could be anticompetitive. That analysis is an essential piece of the puzzle, but it does not address the question of whether the conduct is also consistent with competitive behavior and, if so, how to distinguish among competing explanations for the behavior. Greater attention to competitive behavior will be necessary if economists are to provide law enforcers with what they need. Relatedly, greater attention should be paid to the relative error costs of permitting anticompetitive conduct and of chilling efficient conduct.

Greater attention to competitive behavior will be necessary if economists are to provide law enforcers with what they need.

Another striking feature of the existing literature is how theoretical it is. The appropriate mix of theoretical and empirical investigation is, of course, a complicated issue. Any interpretation of evidence rests, explicitly or implicitly, on some theory or model. It would make no sense, therefore, to suggest that we abandon theory and just look at the facts. Yet, theory unbridled from empirical observation is equally problematic. There are two principal reasons why, we believe, the theoretical nature of the existing literature has limited its usefulness. First, much of the theory about unilateral conduct by a firm with market power is sufficiently abstract that it is hard to know how to match the theory to the facts of any particular situation. We mean this point more as observation than criticism. Understanding some types of behavior at a theoretical level is very complicated. Nonetheless, economists need to acknowledge that much of the existing work remains at too rudimentary a stage to be of practical use. Second, firms without market power sometimes take the same sorts of actions

16 *Microsoft*, *supra* note 1.

17 United States v. Dentsply International, Inc., 399 F.3d 181 (3rd Cir. 2005) [hereinafter *Dentsply*].

18 Timothy J. Brennan, Saving Section 2: Reframing Monopolization Law (AEI-Brookings Joint Center for Regulatory Studies, Related Publication 05-27, Dec. 2005), *available at* http://www.aei.brookings.org/admin/authorpdfs/page.php?id=1202.

that create antitrust problems for firms with dominant shares. The abuse of market power cannot be a reasonable motive in these cases, so an alternative explanation must exist. By itself, this does not justify a conclusion that the effects of the actions when taken by a firm with market power are the same as when they are taken with a firm operating in a competitive market. The baseball player Yogi Berra, who had a penchant for sayings that were simultaneously trivial and wise, famously said, "You can observe a lot just by looking." If economists increase the attention they devote to understanding competitive behavior, observations on firms without market power might be a more fruitful starting point than theoretical models of competition.

In the following section, we make these points—the need to understand competitive behavior, the potential for doing so empirically, and the limits of what we learn from existing theory—with respect to one class[19] of tying behavior, an area that has played prominently in recent cases and where we believe the economics literature has been deeply flawed. Just as formulators of foreign policy must learn the lessons of Vietnam and antitrust practitioners must learn the "lessons of IBM," we believe that antitrust economists should learn the "lessons of tying."

In using this term, we are of course suggesting that for antitrust economists, the tying literature should be viewed as a fiasco on the order of Vietnam or the IBM case. That the literature is problematic might not be evident to all readers, and even if it is, the comparison to Vietnam might seem melodramatic. Yet, when the *Journal of Economic Perspectives* published a symposium on the *Microsoft* case, it contained an article by Michael Whinston, arguably the leading expert on the economics of tying, which said, in essence, that economists do not understand much about tying.[20] We agree with his assessment, and we suggest that antitrust economists should view it as an admission of collective failure from which lessons need to be learned. In Section III, we turn to other types of tying, including bundled discounts. In the wake of *LePage's*,[21] this has become an unsettled aspect of antitrust law. We offer some suggestions for how to avoid the mistakes of the previous tying literature with respect to bundled discounts. In Section IV, we offer suggestions for useful economic analysis related to single-firm conduct more generally. Section V contains some brief concluding comments.

19 Below, we will argue that legal tying doctrine applies to a variety of cases that need to be distinguished.

20 Michael D. Whinston, *Exclusivity and Tying in U.S. v. Microsoft: What We Know and What We Don't Know*, 15 J. ECON. PERSP. 63 (Spring 2001).

21 LePage's Inc. v. 3M, 324 F. 3d 141 (3rd Cir.) (*en banc*), cert. denied, 542 U.S. 953 (2004) [hereinafter *LePage's*].

II. From *Loew's* to *Microsoft* – The Role of Economics

Microsoft[22] was the culmination of the failure of the economics literature on tying. Microsoft's decision to include a web browser in its operating system was central to a case in which the DOJ ultimately sought to break up a firm that, at the time, had the largest market capitalization in the world. We believe that the economics literature was not of great help in clarifying the issues surrounding the tying claim, and it is this failure that creates the need to learn the lessons of tying. The roots of how the economics literature went astray are in *Loew's*.[23] The business practice at issue was the block booking of movies, which in effect tied a studio's B-rated movies (like *Getting Gertie's Garter*) to its A-rated movies (like *Gone with the Wind*). George Stigler's (1968) analysis of the case is widely viewed as a seminal article on bundling and tying.[24] One criticizes Nobel Laureates, particularly one as widely revered as George Stigler, at one's peril. Yet, in assessing where a body of literature went awry, we need to assess the first, seminal steps. Stigler cannot be blamed for the failure of others to correct the course, but the problems that have limited the usefulness of the tying literature are evident in his analysis.

A. *LOEW'S* AND THE ECONOMIC ANALYSIS OF BUNDLING

Prior to Stigler's analysis, the presumption was that tying was a way of leveraging market power from one good to another good for which the market was inherently more competitive. Stigler questioned the conventional wisdom about leveraging and posed an alternative explanation based on a simple numerical example that is worth repeating here. He hypothesized two movies (X and Y) to be licensed to two movie theaters (I and II). Theater I was willing to pay $8,000 for X and $2,500 for Y. Theater II was willing to pay $7,000 for X and $3,000 for Y. Stigler observed that with simple pricing, the distributor would have to charge $7,000 for X and $2,500 for Y, yielding total revenue of $9,500 per theater. By tying the two together, however, it could charge $10,000 for the package.[25]

Stigler's analysis served a useful purpose by posing a fundamental question: What did tying accomplish that could not be accomplished with simple monopoly pricing? Any theory of monopoly leveraging must come to grips with this very basic question. Still, even making due allowance for the proposition that pathbreaking analyses require elaboration and refinement, there is much to crit-

22 *Microsoft*, *supra* note 1.

23 United States v. Loew's Inc., 371 U.S. 38 (1962) [hereinafter *Loew's*].

24 George J. Stigler, *A Note on Block Booking, in* The Organization of Industry 165 (1968).

25 *Id.* at 165-170.

icize in Stigler's analysis of this problem. The role of economic theory in a problem such as this one is to strip away inessential details to bring the most important features into sharp relief. For a variety of reasons, Stigler's analysis did not get to the heart of the matter.

First, the analysis ignored the distinction between tying and bundling. Bundling is the sale of two goods in combination that could be sold separately. By itself, selling a bundled product does not preclude selling the components separately—the practice now known as mixed bundling. The issue in *Loew's* was tying, not bundling. While the result in Stigler's example was that tying would occur, the underlying assumptions were not rich enough[26] for mixed bundling to be a very interesting strategy.

Second, the link between the assumptions of the example and the facts of the case were tenuous. The example is, at best, a logically possible explanation for why firms might tie two goods that could be sold separately rather than a compelling explanation for why movie distributors tied the particular movies that they did.

Third, the explanation was based entirely on demand and not at all on costs. This problem is related to the previous one, as Stigler's explanation showed little appreciation for the economics of film production, film distribution, and film exhibition and the relationship among the different stages.

B. TYING AND *MICROSOFT*

Subsequent developments in the literature on bundling and tying addressed some but not all of these problems. The distinction between mixed bundling and pure bundling emerged relatively quickly.[27] As a fairly general proposition, mixed bundling generates higher profits than pure bundling if bundling does not affect costs.[28] To be sure, the optimal prices under mixed bundling can entail charging a premium for the bundle. If nothing prevents customers from buying all the components separately, then mixed bundling might not be a feasible strategy. One might imagine a theory of tying based on those cases where the optimal mixed bundle would entail a premium for the bundle, although it is not clear that pure bundling would dominate selling the components separately in these cases.

26 The problem is that with mixed bundling, there are three distinct products; and any interesting theory of mixed bundling would predict positive demand for all three. With only two customers in the model, there cannot be three distinct buying patterns.

27 Walter J. Adams & Janet L. Yellen, *Commodity Bundling and the Burden of Monopoly*, 90 Q. J. ECON. 475 (1976). Mixed bundling is the practice of selling the goods separately and in bundled form with a discount for the bundle. Pure bundling means that the firm sells the goods only in bundled form.

28 R. Preston McAffee, John McMillan & Michael D. Whinston, *Multiproduct Monopoly, Commodity Bundling, and Correlation of Values*, 104 Q. J. ECON. 371 (1989).

Michael Whinston (1990) put forward the first model generally accepted by economists as a logically consistent theory of anticompetitive tying when prices are not regulated.[29] He posited a two-product firm with a monopoly over one of the goods it sells. In the other good, which is produced with increasing returns to scale, it faces potential competition from an entrant. Whinston showed that by tying the two goods together, the two-product firm could, as a matter of theory, deny the entrant adequate scale and keep it out. The article was fundamental in pointing out the limitations of the single-monopoly profit theorem. Yet, as Whinston was careful to point out, the practical implications of the model were not clear. In the tradition of modern industrial organization theory, Whinston's article laid bare the broad outlines of a logically sound case that tying could be anticompetitive. Yet, like much basic research, it left for others a great deal of development work to flesh out how to apply the model in practice.[30]

One of the striking features of the *Microsoft* case is the prominent role that tying played. This was likely the result of legal doctrine, not economic analysis. For a firm with market power in the tying good, tying remains a per se violation.[31] One cannot know for sure whether the government would have presented its case differently if the legal standards for tying were more similar to related practices, but it seems plausible that it would. The key question economists should be asking about the tying claims in *Microsoft* is whether economists successfully laid out the economics of the tying claims.

One might argue that they did. Many discussions of tying point out that tying is a common occurrence that, in most instances, lowers costs or provides convenience.[32] Moreover, even the new theory sometimes taken as support for the case, such as that found in Carlton and Waldman (2002),[33] suggests an implicit recognition of the need to distinguish the specific tying at issue in the case from most tying. The nature of this analysis was to extend the basic logic of Whinston's article[34] to assumptions that more nearly resembled the setting of the case.

29 Michael D. Whinston, *Tying, Foreclosure, and Exclusion*, 80 Am. Econ. Rev. 837 (1990).

30 For a more detailed discussion of the Whinston models and its practical limiatations, *see* Keith N. Hylton & Michael Salinger, *Tying Law and Policy: A Decision-Theoretic Approach*, 69 Antitrust L.J. 469 (2001).

31 Jefferson Parish Hospital District No. 2 et al. v. Hyde, 466 U.S. 2 (1984).

32 Barry Nalebuff, Bundling, Tying, and Portfolio Effects 46 (DTI Economics Paper No. 1, Part 1 – Conceptual Issues, Feb. 2003).

33 Dennis W. Carlton & Michael Waldman, *The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries*, 33 Rand J. Econ. 194 (2002).

34 Whinston, *supra* note 29.

We believe that this assessment of the economics literature paints far too flattering a picture. The role of economic analysis in antitrust (and in general) is to focus on the most important aspects of a problem. In this regard, we believe that the economics literature on tying missed essential elements of the sort of tying in question in *Microsoft*. Even when the economics literature has acknowledged efficiencies from tying, it has failed to make the fundamental distinction between efficiencies of bundling and efficiencies of tying. Efficiencies of bundling are cost savings or convenience that arise from providing a bundled product to people who want all the components. The prototypical example is shoes. Virtually everyone who wants shoes wants them in pairs,[35] and it is obviously cheaper to provide them in pairs than it is to sell them separately. Efficiencies of tying are more subtle. One must consider why it is efficient **not** to provide the individual components to those who want just some of them rather than ask why it is efficient to provide a bundle to people who want all the components. The fact that shoes are chosen as the example to illustrate the efficiencies of bundling/tying suggests that this distinction has not been adequately appreciated. Virtually no one wants to buy right and left shoes separately. One can, of course, pose the question of why someone who lost his right shoe could not replace it without buying an entire pair. In the context of that example, the question sounds more philosophical than practical.

ONE MUST CONSIDER WHY IT IS EFFICIENT **NOT** TO PROVIDE THE INDIVIDUAL COMPONENTS TO THOSE WHO WANT JUST SOME OF THEM RATHER THAN ASK WHY IT IS EFFICIENT TO PROVIDE A BUNDLE TO PEOPLE WHO WANT ALL THE COMPONENTS.

To understand the efficiencies of tying, one must recognize that tying represents a choice to offer a subset of the products that a firm could conceivably offer. An efficiency explanation for such a decision must rest on a cost of product offerings. There is an old literature in economics on product selection that poses the question of whether the set of products offered by the market is the optimal set.[36] To get at the essence of that problem, it is standard in that literature to assume a fixed cost of each product offering. We are not aware of anyone suggesting that the product selection literature and the assumptions underlying it are relevant for tying analysis until recently.

35 Mark Frankena has pointed out to us that there is indeed a market for single shoes. Yet, with the exception of Birkenstock, from which individual shoes can be ordered specially, the rest of the market seems to be from resellers that untie the tied offerings generally available. *See Birkenstock Express special orders for extended sizes and single shoes and sandals*, *at* http://www.birkenstockexpress.com/Services/specialorders.cfm/topnav2.256 (last visited Jan. 30, 2006) and *NLLIC ACA Fact Sheet: Mismatched and Single Shoes*, *at* http://www.amputee-coalition.org/fact_sheets/oddshoe.html (last visited Jan. 30, 2006).

36 Avinash K. Dixit & Joseph E. Stiglitz, *Monopolistic Competition and Optimum Product Diversity.*, 67 Am. Econ. Rev. 297 (1977) and Michael A. Spence, *Product Selection, Fixed Costs, and Monopolistic Competition* , 43 Rev. Econ. Stud. 217 (1976).

Even Carlton and Waldman[37] address only the issue of whether the tying of the browser to the operating systems could, as a theoretical matter, be monopoly extention. The article does not address the question of how to tell whether the tie was a response to competitive pressure rather than an effort to thwart it. If we are to distinguish anticompetitive tying from competitive tying, we need to understand competitive tying better than we do. Observing tying under competition would seem to be a productive way to start.

In joint work with David Evans (2005),[38] Salinger has made much of the example that electrical plug adapters are sold at a variety of outlets only in packages of four different adapters, effectively forcing people to buy adapters they do not want to obtain the one that they do.[39] In terms of the total amount of commerce involved, the example is trivial. Yet, the economics literature on tying did not provide a compelling explanation for why the adapters were not sold separately. Certainly, the case seemed different from shoes. By focusing on the adapter, we developed a theory of competitive tying that went beyond the casual observations in the literature that of course much tying is to save costs or provide convenience. In particular, the theory led to two key insights. First, to understand competitive tying, one needs to understand the scale economies associated with each individual product offering, recognizing that a firm selling a bundle of two products as well as the two products separately is selling three distinct products. Without a fixed cost (or, more generally, a scale economy) associated with each product offering, a competitive firm's refusal to sell components separately to those who do not want all the components of a bundle makes no sense. Second, once one recognizes the fixed costs associated with individual product offerings, tying can arise under competition in circumstances that had not been acknowledged previously (at least in the formal economics literature). There had been a presumption that competitive tying would occur when everyone (or virtually everyone) wanted all the components. When fixed costs of product offerings are taken into account, however, tying can occur even if no one wants all the components. Companies might tie in order to meet the needs of diverse customers with a single product.

Once this possibility is pointed out, it is obvious that such tying is common. No one reads all of the morning newspaper and few households watch every channel they receive as part of their cable television package. If one objects that these services are not competitively supplied, plenty of competitive examples

37 Carlton & Waldman, *supra* note 33.

38 David S. Evans & Michael Salinger, *Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law*, 22 YALE J. ON REG. 37 (2005).

39 Originally, we focused on the package available at RadioShack. Interestingly, it now offers a package that eliminates (or at least substantially reduces) the extent of tying. RadioShack sold the package of four adapters without selling all the individual adapters separately for at least several years, and the tying was not limited to RadioShack.

exist as well. Few university students take advantage of every service that tuition entitles them to. No one rides every ride at amusement parks that charge a lump sum admission rather charging for each ride.

At the trial in *Microsoft*, there was testimony from customers who would have preferred that Microsoft's operating system, Windows, not include its web browser application, Internet Explorer.[40] While such testimony might have been useful to satisfy the legal standard that the operating system and browser were separate products, such testimony could not possibly have done much to suggest that the tie was anticompetitive. Such a conclusion would rest on the presumption that when tying occurs in competitive markets, consumers do not end up purchasing components or product features that they would prefer to do without. That simply is not the case.

We believe that the adapter example led to important insights about tying. Perhaps that claim is immodest (for one of us), but we risk that appearance to make a more general point about the literature. Simple observation could have such high marginal value because there has been too little effort devoted to observation (and, we would suggest, too much devoted to theory). This, in our view, is one of the most important lessons of tying, and it is a lesson that must be learned with respect to other aspects of single-firm conduct if we are to avoid similar problems in our analysis of those practices.

III. *LePage's* and Bundled Discounts

A case of considerable recent interest is *LePage's, Inc. v. 3M*.[41] The defendant, 3M—the manufacturer of Scotch Tape—had a dominant market share in the market for transparent tape and sold other products as well. The plaintiff, LePage's, was the principal alternative supplier of private label transparent tapes. One of 3M's other products was private label tape that competed with LePage's offerings. At issue in the case was 3M's pricing practices whereby customers could obtain discounts on 3M products based on purchase volume and growth targets on a wide range of 3M products. Failure to meet the target in any one product line resulted in forfeiture of the entire rebate, irrespective of sales and growth in other product lines. Since one of the qualifying product lines was 3M's private label adhesive tape, the price retailers paid for 3M's branded tape depended on the quantity of 3M's private label tape that it purchased. LePage's attributed its decline in share of the private label "market" (from 88 percent to 67 percent) to 3M's bundled discount.[42]

40 United States v. Microsoft Corp. 87 F. Supp.2d 30, 48-49 (D.D.C. 2000).

41 *LePage's*, *supra* note 21.

42 Sales of private label transparent tape comprise only a small portion of the total U.S. transparent tape market.

LePage's alleged that 3M's bundled rebate program constituted unlawful monopolization in violation of Section 2 of the Sherman Act. A jury found for LePage's on its Section 2 claim and awarded nearly US$23 million in damages (US$69 million trebled).[43] A three-judge panel of the U.S. Court of Appeals for the Third Circuit reversed with one judge dissenting. Subsequently, after an en banc rehearing, the Third Circuit vacated that decision and affirmed the district court with the author of the earlier panel decision now dissenting.

The antitrust perspective on bundled discounts remains a matter of dispute. Relying on *Brooke Group*,[44] 3M argued that as a matter of law, above-cost pricing, no matter what its exclusionary effect, cannot make out a claim under the Sherman Act. In this case, although 3M may have forsaken some short-term profits as a result of its awards of rebates, at no time were its sales unprofitable (i.e., below cost). Therefore, according to 3M, there could be no proof of injury to competition.

The court rejected this argument and found that, even if above-cost pricing is not generally unlawful, *Brooke Group* applies only where the claim is predatory pricing by a monopolist.[45] In this case, according to the court, the challenged conduct was exclusionary irrespective of its not being predatory. Citing Areeda and Hovenkamp for authority, the court chose to analogize 3M's bundled rebate programs with "tying" where the anticompetitive effects are in the form of foreclosure of rivals. In so doing, the court eschewed a pure pricing analysis, and instead relied on its earlier opinion in *SmithKline Corp. v. Eli Lilly & Co.*[46] In both cases, the court found that the defendants' bundled rebates reflected an exploitation of (otherwise legal) monopoly power by linking a product that faced competition to a product that did not.

The dissenting judge argued that, absent a showing of below-cost pricing, the evidence must show some other basis for the Section 2 violation and that the evidence in *LePage's* did not do this. Among the factual matters he deemed relevant were plausible business justifications for the rebate programs based on distributional efficiency.

This is precisely the sort of issue where clarification from economists could be of help. Just as the courts have struggled with how to view bundled discounts,

43 U.S. antitrust law allows private parties to sue for treble damages.

44 *Brooke Group*, *supra* note 9.

45 *LePage's*, *supra* note 21, at 151-52.

46 SmithKline Corp. v. Eli Lilly & Co., 575 F.2d 1056 (3d Cir. 1978).

economists have debated the issue.[47] Without prejudging the answer, let us offer some suggestions on fruitful lines of inquiry that will reduce the chances that the bundled discount literature will suffer from the same problems as the tying literature. First, analogies to other kinds of practices are of some—but limited—use. The practice is related to predatory pricing in that some way is needed to distinguish the bundled discounts that might be objectionable from non-linear pricing schemes that are not. A natural cut-off to consider is that when the entire discount is applied to one of the goods, then the price for the good is below incremental cost. For the remainder of this paper, we refer to this situation as "extreme bundled discounts." The similarity of that cut-off to standards for predatory pricing does not, however, mean that such cases should be tried under predatory pricing doctrine. In particular, if one could show that the practice is inherently profitable, as may be possible,[48] then it should not be necessary to demonstrate the plausibility of subsequent recoupment, as is the case in a standard predatory pricing claim.

Similarly, the observation that the practice is like tying is both true and of only limited help. Literal tying requires that a company refuse to sell one product without another. Even when companies do not literally tie, their pricing can create a virtual tie if buying one product without another is not an economically viable option. Extreme bundled discounts can be virtual ties. While, from a purely legal standpoint, that might be useful for understanding how current law might be applied to such a case, it is of little use for clarifying what antitrust policy toward extreme bundled discounts should be. The problem is that legal tying doctrine applies to a wide variety of cases that should be viewed as economically distinct.

For example, the analysis of tying behavior discussed in Section II above applies to cases when the seller charges a single combined price for two or more goods. The essence of the decision is not to charge separately for one and perhaps all the component goods. Another class of tying concerns systems consisting of a durable (like a camera) and a consumable (like film).[49] A common term, "tying," is used for both cases, but the practices are much different. In durable/consumable cases, the two goods are sold (and charged for) as distinct items, so tying does not save the cost of one or more product offerings. There might well be efficiency rationales for tying in these types of cases, such as the

47 Patrick Greenlee, David Reitman, & David S. Sibley, An Antitrust Analysis of Bundled Loyalty Discounts (Economic Analysis Group, Working Paper No. EAG 04-13, Oct. 2004), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=600799.

48 With bundled discounts, the undiscounted price might not be one that the seller wishes to charge. Thus, the discount cannot be considered a profit sacrifice the same way prices below marginal cost can be in standard predatory pricing.

49 Michael Salinger, *Business Justification Defenses in Tying Arrangements (2005)*, *in* Issues in Competition Law and Policy (Wayne D. Collins ed., forthcoming 2006).

desire to avoid assessing blame for the failure of a system comprised of parts sold by different suppliers, but those efficiencies are different from those in the adapter (or newspaper) examples.

In at least one way, bundled discounts are more similar to this latter type of tying in that the products for which the bundled discount applies are priced as separate products. Retailers are able to vary the amount of branded and unbranded tape they purchase, and they pay distinct (discounted) prices for each. That said, neither the plausible efficiency nor the pricing consequences of bundled discounts are similar to the durable/consumable case.

Game theoretic analyses of bundled discounts, such as that in Greenlee and Reitman (2004),[50] help somewhat, but they may be of more interest to economists than to law enforcers. If, in a case like *LePage's*, the marginal price for one of the goods is below variable cost, we expect that courts will find the pratice inherently suspect regardless of whether economists can justify the practice as a Nash equilibrium strategy.

What would be far more useful to law enforcers is evidence of bundled discounts by firms operating in competitive markets, if indeed such evidence exists. If the practice is more widespread than is commonly believed, that should not make the practice per se legal. Observations of the practice under competitive conditions should, however, provide an opportunity to understand any efficiency motivations. As extreme bundled discounts typically arise in sales by manufacturers to retailers, they are harder to observe than the tying of products sold to final consumers. Thus, economists will need cooperation from firms that have such pricing policies to get the information they need to provide the analysis policymakers want. To the extent that firms want to engage in extreme bundled discounts and want to make the case in the upcoming Section 2 hearings that antitrust law should not be hostile to them, they might consider working with economists to help sort out what should be viewed as efficiencies from the standpoint of public policy rather than private efficiencies (which might indeed be anticompetitive).

WHAT WOULD BE FAR MORE USEFUL TO LAW ENFORCERS IS EVIDENCE OF BUNDLED DISCOUNTS BY FIRMS OPERATING IN COMPETITIVE MARKETS, IF INDEED SUCH EVIDENCE EXISTS.

If evidence of efficiency motivations for extreme bundled discounts does exist, then the sort of game theoretic analysis that has become the starting point for economists will be of value. To be useful, however, such models must go beyond the demonstration of the possibility of anticompetive bundled discounts. The models will have to generate insights about what observable factors can be used

50 PATRICK GREENLEE & DAVID REITMAN, COMPETITING WITH LOYALTY DISCOUNTS (Economic Analysis Group, Working Paper No. EAG 04-2, Feb. 2004), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=600799.

to disinguish anticompetitive bundled discounts from those that are efficient and/or be formulated in a way that they give rise to serious empirical estimates of the cost of letting anticompetitive instances of the practice go unchallenged.

IV. Useful Economic Analysis of Unilateral Conduct More Generally

The upcoming Section II hearings will provide an opportunity to comment on a wide range of single-firm conduct and for economists to provide useful analysis. In addition to tying, bundling, and bundled discounts, topics will likely include exclusive dealing, loyalty discounts, refusals to deal, and full-line forcing, to name a few. To avoid repeating the mistakes of tying, we offer some suggestions for what type of economic analysis will be useful:

- First, it is important not to accept legal categories as being economically relevant. We made this point with respect to tying, in which there are important sub-classes that need to be distinguished.

- Second, in principle, the appropriate simple legal rules will depend on the relative frequency of competitive and anticompetitive instances of the practice in question. While there is no practical way to get objective estimates of the proportions, careful observation of the practice under competition will be informative. This will serve two purposes. First, while arguably not scientific, the ease of finding examples of the practice under competition is a reasonable factor to consider in forming subjective assessments. Second, once examples are found, exploring the rationale for the practice can lead to an understanding of the nature of the efficiencies that was not previously obvious.

- Third, while admittedly difficult, documenting cases in which the practices were in fact anticompetitive will be extremely useful. Elsewhere, Salinger (2005) has argued that documented instances of anticompetitive tying are rare and may not exist.[51] Some of our current and former colleagues from the FTC have argued this point more generally with respect to vertical restraints in general.[52] Documented cases will do far more to justify antitrust hostility to the practices than mere theorizing.

None of this is to suggest that theory cannot be helpful. But, it is not the place to start.

51 Salinger, *supra* note 49. *See also*, Hylton & Salinger, *supra* note 30.

52 James Cooper et al., *Vertical Antitrust Policy as a Problem of Inference*, 23 Int'l J. Indus. Org. 639 (2005).

V. Concluding Comments

Economists have, and will continue to have, opportunities to provide insights that lead to better legal rules. These opportunities will be lost, however, if the recommendations fail to address the question of how to distinguish among competing explanations for the practices at issue and if they are based on theory that is too abstract to match real settings.

Volume 2 | Number 1 | Spring 2006

Competition Policy International

The Law of Exclusionary Pricing

Herbert Hovenkamp

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

The Law of Exclusionary Pricing

Herbert Hovenkamp

The success of the Areeda-Turner test for predatory pricing and the U.S. Supreme Court's adoption of demanding proof requirements in its 1993 *Brooke Group* decision have made it very difficult for plaintiffs to win conventional predatory pricing claims. While many challenges to exclusionary pricing continue to be made, the legal theory has evolved away from classical predation to a variety of other theories. These theories include challenges to quantity and market share discounts, single item and package discounts, and various purchasing practices, including slotting fees, overinvestment in fixed cost assets, and overbuying of variable cost inputs. Plaintiffs have enjoyed somewhat greater success with these alternative theories, in large part because the practices are not as well understood as conventional price predation is. This paper examines the state of the law of both conventional predatory pricing and these more recent variants and offers some recommendations.

The author is the Ben V. & Dorothy Willie Professor of Law and History at the University of Iowa, College of Law.

I. Introduction

The thirty years since the publication of Areeda and Turner's landmark article on predatory pricing have witnessed a revolution in the antitrust law of exclusionary pricing.[1] The result has been that classical predatory pricing complaints have nearly disappeared from the antitrust case law, and plaintiffs rarely win them.[2] Nevertheless, the law of exclusionary pricing has hardly disappeared. Rather it has morphed into the law of anticompetitive discounts, slotting fees, or spending.

Areeda and Turner's 1975 article, subsequently expanded in Volume 3 of the *Antitrust Law* treatise,[3] observed that predatory pricing was rational only if the predator could recoup its investment in predation with a comfortable period of post-predation monopoly profits.[4] In addition they argued, because the danger of false positives is considerable, predatory pricing should be condemned only on prices that are below cost, and that the most useful measure of cost is either short-run marginal cost (MC) or average variable cost (AVC). Areeda and Turner added that in most cases AVC is the better measure because it is typically easier to compute in litigation.[5]

While the U.S. Supreme Court has never passed judgment on the correct price/cost test for predatory pricing, the U.S. Circuit Courts have generally agreed that either marginal cost or average variable cost is the correct number. Only the U.S. Court of Appeals for the Eleventh Circuit adheres to an average total cost

1 *See* Phillip E. Areeda & Donald F. Turner, *Predatory Pricing and Related Practices Under Section 2 of the Sherman Act*, 88 HARV. L. REV. 697 (1975) (*reprinted in* 1 COMPETITION POL'Y INT'L 2, 177-212).

2 One important, recent counterexample is Spirit Airlines, Inc. v. Northwest Airlines, Inc., 431 F.3d 917 (6th Cir. 2005) (denying summary judgment on small air carrier's predatory pricing claim against Northwest Airlines) [hereinafter *Spirit Airlines*].

3 3 PHILLIP E. AREEDA & DONALD F. TURNER, ANTITRUST LAW, at Ch. 7C (1978); the current version is 3 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW, at Ch. 7C (2d ed. 2002 & 2005 Supp.).

4 *See* Areeda & Turner, *supra* note 1, at 698:

> [P]redation in any meaningful sense cannot exist unless there is a temporary sacrifice of net revenues in the expectation of greater future gains. Indeed, the classically-feared case of predation has been the deliberate sacrifice of present revenues for the purpose of driving rivals out of the market and then *recouping* the losses through higher profits earned in the absence of competition. Thus, predatory pricing would make little economic sense to a potential predator unless he had (1) greater financial staying power than his rivals, and (2) a very substantial prospect that the losses he incurs in the predatory campaign will be exceeded by the profits to be earned after his rivals have been destroyed. (emphasis added)

These concerns were restated in 3 ANTITRUST LAW ¶ 711b (1st ed. 1978).

5 *See* 3 ANTITRUST LAW ¶ 740 (2d ed. 2002).

test.[6] While Areeda and Turner never elaborated very much on their recoupment requirement, other than to state that it required a close appraisal of monopoly power and entry barriers, subsequent literature did. In its 1993 *Brooke Group* decision, the Supreme Court assessed a stringent recoupment requirement.[7]

These twin requirements have proven to be devastating for most predatory pricing plaintiffs in the federal courts. The requirements are well-established in the case law and unlikely to be rejected anytime soon. Taken together, they almost certainly make the law of predatory pricing somewhat underdeterrent; that is, current law probably fails to recognize some instances of exclusionary pricing whose overall welfare effects are negative. But it is equally clear that a more lenient set of rules would produce many challenges and chill a great deal of aggressive, but pro-competitive, pricing. The social cost of the resulting limits on competition would almost certainly be much greater than that of the occasional instance of anticompetitive strategic pricing that goes unrecognized.

The antitrust law of predatory pricing, more than any other area, is dedicated to the principle that the social cost of false positives in antitrust analysis is higher than the cost of false negatives.[8] False positives often will induce firms not to price aggressively for fear of large treble damage awards. As a result, their impact reaches far beyond the parties to a particular lawsuit and can cause significant harm to the economy. By contrast, false negatives are thought to be much rarer and, as a result, they affect only a few firms in a few situations. The number of markets that are structurally conducive to durable monopoly created by predatory pricing is undoubtedly quite small. Further, the natural forces of competition are more likely to correct for false negatives. To be sure, these propositions are difficult to test, but they seem intuitively correct.

Notwithstanding the numerous criticisms to the MC/AVC test that have been addressed, no alternative has proven to be more reliable or more workable. The intuition behind using reasonably anticipated marginal cost[9] is that competition drives prices to that level, and firms cannot profit when they go below it. As a result, prices above marginal cost are consistent with at least competitive returns, and prices below marginal cost require an explanation. The intuition behind

6 *See* McGahee v. Northern Propane Gas Co., 858 F.2d 1487, 1500 (11th Cir. 1988), *cert. denied*, 490 U.S. 1084 (1989).

7 Brooke Group Ltd v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993) [hereinafter *Brooke Group*]; *see* 3 ANTITRUST LAW ¶ 726 (2d ed. 2002).

8 *See* HERBERT HOVENKAMP, THE ANTITRUST ENTERPRISE: PRINCIPLE AND EXECUTION 158-174 (2006).

9 We speak of "reasonably anticipated" MC or AVC because the number must be measured ex ante. Whenever there is a time lag between production and sale, a firm must guess at how much to produce and may have to estimate both input costs and market price. Firms cannot be penalized because ex post prices fall below the cost measure where the firm's reasonable expectations were to the contrary.

using AVC as a surrogate is that, first, in a competitive equilibrium MC and AVC are very close to each other. Second, AVC is theoretically easier to measure—one simply identifies the firm's variable costs over a defined time period, and divides this number by the number of units of output. Indeed, AVC is more than a surrogate. It is independently defensible because AVC is the typical firm's "shutdown" price. That is, a firm will continue in production as long as it is recovering variable costs and making some contribution to fixed costs; but, once prices fall below AVC, production itself is costly. To be sure, practical problems are considerable. The line between variable and fixed costs is often ambiguous and joint costs are very difficult to take into account. Nevertheless, AVC is typically easier to measure than MC.[10]

As legal tests, both marginal cost and average variable cost are fairly crude attempts to equate non-predatory price levels with sustainability. Prices below MC or AVC are non-sustainable because the firm has greater costs than revenues and must eventually exit from the market. A few writers have equated the recoupment requirement with lack of sustainability. In that view, the *Brooke Group* showing of recoupment is necessary only if the law requires prices below cost, because when prices are above cost there is nothing to recoup.[11] But this position ignores the fact that predatory pricing strategies are costly to the firm whether or not the predatory price is below or above cost. Consider the dominant firm that has costs of 6 and a short-run profit-maximizing price of 10[12] but who expects entry to the competitive level to occur within four years if it charges that price. However, the firm also calculates that a price of 8, still well above its costs, would deter new entry indefinitely. While the price of 8 is completely sustainable, this hardly entails that the firm who employs this strategy has nothing to recoup. During the first four years it will earn less, and this loss must be regarded as an investment in the longer stream of supracompetitive profits that it anticipates. The investment is profitable only if the longer stream of sales at a price of 8 generates greater profits than the shorter stream of sales at a price of 10. In sum, even if predatory pricing law abandoned the requirement of below-cost sales, some theory of recoupment would very likely be required—indeed, the requirement would be even more important because the risk of false positives is significantly higher when courts are authorized to condemn above cost prices as predatory.

EVEN IF PREDATORY PRICING LAW ABANDONED THE REQUIREMENT OF BELOW-COST SALES, SOME THEORY OF RECOUPMENT WOULD VERY LIKELY BE REQUIRED.

10 *See* 3 Antitrust Law ¶¶ 740, 742 (2d ed. 2002).

11 *E.g.*, Aaron S. Edlin, *Stopping Above-Cost Predatory Pricing*, 111 Yale L.J. 941, 942 (2002); Einer Elhauge, *Why Above-Cost Price Cuts To Drive Out Entrants Are Not Predatory—and the Implications for Defining Costs and Market Power*, 112 Yale L.J. 681, 697 (2003).

12 A firm's short-run profit-maximizing price is the price determined by equating the firm's immediate marginal costs and marginal revenues without considering the impact of this price on expansion by or entry of rivals.

One problem with using AVC rather than MC as a measure of predation is that at high levels of output, where predation presumably occurs, MC and AVC diverge, with MC higher than AVC.[13] Assuming that MC is the theoretically correct measure, the AVC surrogate becomes increasingly favorable to defendants as output increases beyond the plant's optimal level. This result has prompted some critics to label the AVC test a "defendant's paradise."[14] Of course, this is fully consistent with the observation that the current law of predatory pricing is somewhat underdeterrent, but that false negatives are not nearly as damaging in this situation as false positives.

Another problem with the marginal and average variable cost tests has been the tendency to measure these costs too myopically, considering only the shortest possible run. In some markets, such as those having significant intellectual property components, short-run marginal cost would produce prices that are much, much lower than the sustainable level. For example, consider computer software, where development costs might run into the hundreds of millions of dollars, but short-run production costs consist of little more than the cost of stamping a CD-ROM and packaging, or become virtually zero in the case of downloadable software. In such cases, a sustainable price must be sufficient to amortize the firm's R&D development. In order to be effective, a predatory pricing test would have to attribute some element of R&D costs to each unit of production. But doing this is extraordinarily difficult unless the product has already exhausted its commercial life, given that most firms set price without knowing how many units they will sell over the product's lifecycle. A firm that anticipated sales of 1,000,000 units might later be charged with predatory pricing if it ended up selling only 300,000, thus entailing larger per unit production costs. To date, these problems have no administrable solutions.

A good example of myopia in the computation of AVC is the *American Airlines* case, where the court refused to consider opportunity costs in determining whether American had charged below-cost prices.[15] In response to entry by small carriers American not only cut prices drastically, it also shifted aircraft from profitable routes elsewhere in order to flood the routes where competitive entry had occurred. The court refused to consider the revenue foregone from the vacated routes as part of the cost of predation. The court incorrectly characterized the government's theory as showing, not that prices were below cost, but that American was simply not maximizing its profits.

13 *See* 3 ANTITRUST LAW ¶ 740c (2d ed. 2002).

14 *See* Oliver E. Williamson, *Predatory Pricing: A Strategic and Welfare Analysis*, 87 YALE L.J. 284, 305 (1977).

15 United States v. AMR Corp., 140 F.Supp.2d 1141 (D.Kan. 2001), *aff'd*, 335 F.3d 1109 (10th Cir. 2003).

To be sure, excessive speculation about opportunity costs could lead courts on fishing expeditions for ways that a defendant might have earned more by selling a different product or selling it in a different place. But that was not what the government was asking for in this case. Known aircraft were being shifted into the predatory routes and shifted back again once rivals had been driven from the market. The U.S. Court of Appeals for the Sixth Circuit appraised opportunity costs much more realistically in the *Spirit Airlines* decision.[16]

As for the recoupment requirement, its real impact has been to require courts to take two factors much more seriously. The first is the cost of a predation scheme, focusing particularly on its duration. The second is the responsiveness of actual and potential rivals, which necessitates inquiries into the height of entry barriers, the disposition of victims' assets,[17] and the ability of existing rivals to increase output quickly in response to higher prices.

As presently formulated, the recoupment requirement makes sense when firms are attempting to create a monopoly or leverage up a dominant position by cutting prices. Predation is costly. The more costly it is, the greater the payoff must be if predation is to be a rational investment. Given our very poor abilities to identify predatory pricing strategies, the recoupment requirement serves to limit predatory pricing liability to those cases where predatory pricing is not a self-deterring strategy. If recoupment is not in the cards, judicial intervention, with its propensity to error, is unnecessary.

Ironically, the poorest case for insisting on a strong recoupment requirement is *Brooke Group* itself, where the U.S. Supreme Court developed the concept. *Brooke Group* did not involve monopoly predatory pricing, but rather predation that was intended to discipline a fairly durable oligopoly that had shown some signs of instability. An important difference between monopoly and oligopoly predatory pricing is that in the monopoly case the predator is bent mainly on destroying its victims, while in the oligopoly case it is intent mainly on bringing them back into the fold. As a result, the alternatives facing the victim of oligopoly pricing are much more attractive than those facing the victim of a monopolization scheme. In the oligopoly case, the firm can either face predatory losses (or returns that are no better than competitive), or else it can rejoin the oligopoly equilibrium and earn high profits. For this reason, disciplinary pricing in an oligopoly is much more likely to be a rational strategy than it is in the monopoly situation. In that case, an overly lenient predatory pricing rule can serve to sta-

16 *See supra* note 2.

17 For example, sometimes victims' assets are auctioned off to competitors in bankruptcy proceedings at very low prices, giving the competitors increased capacity and a lower fixed cost base. *See* 3 ANTITRUST LAW ¶ 729f (2d ed. 2002); and *see* Cargill v. Monfort, 479 U.S. 104, 119 n. 15 (1986), which noted the problem.

bilize oligopolies. Once again, coming up with an administrable rule is extraordinarily difficult. We certainly do not want to condemn price cuts to the competitive level even though these may be all that is necessary to discipline a maverick in an oligopoly market.[18] In the *Brooke Group* case, however, prices were cut to levels well below variable costs.

II. The New Frontiers of Price Predation Claims

The new antitrust challenges to unilateral pricing practices have focused on strategies that are perhaps best characterized as purchases of exclusionary rights. They are sometimes referred to as quasi-exclusive dealing, or quasi-tying. In a quasi-exclusive dealing situation, the dominant firm might offer a lower price in exchange for a purchaser's agreement to:

(a) purchase all of the covered goods from the defendant;

(b) purchase a specified quantity from the defendant; or

(c) purchase a specified minimum share of its total purchases from the defendant.

In the quasi-tying practices, the defendant might:

(a) offer a discount in exchange for an agreement to purchase two products jointly; or

(b) offer a discount that is aggregated across multiple products, typically by pegging the discount to gross sales of a list of products, rather than on each product individually.

In yet another scenario, a supplier makes up front payments to a retailer for exclusive access to a specified amount of shelf space—so-called "slotting" fees.

The migration in the case law from older, head-on challenges to single-product prices as predatory is not difficult to understand. Given the general lack of success experienced by post-*Brooke* predatory pricing plaintiffs, a new approach was needed. Some of the new challenges take advantage of the fact that certain vertical practices, particularly tying, have been treated under more aggressive legal tests than have been applied to simple price cutting. Tying is still nominally covered by an aggressive, but misconceived, per se rule.[19] Exclusive dealing is addressed under the rule of reason and proving illegality is difficult enough for

18 *See* Ian Ayres, *How Cartels Punish: A Structural Theory of Self-Enforcing Collusion*, 87 COL. L. REV. 295, 302 (1987).

19 On this odd per se rule, *see* 9 PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶ 1720 (2d ed. 2004).

plaintiffs; however, it does not require a showing of prices below cost or recoupment, and plaintiffs do continue to win a few cases.[20]

But are these various discounting practices sufficiently unlike conventional predatory pricing to warrant departure from *Brooke Group* standards? Traditional tying and exclusive dealing are typically long-term contractual arrangements or offerings. The buyer can purchase the good subject to the exclusive agreement only by breaching its contract or else by giving up something else in which it has made a significant investment. For example, a franchise tying or exclusive dealing agreement typically requires the franchisee or dealer to purchase the supplier's good exclusively. The buyer can purchase the good from rivals only by giving up its franchise or dealership, which may be far more valuable to the dealer than the value of any savings from an alternative purchase, particularly if the dealer has significant sunk costs invested in its dealership.[21] The result is that an equally efficient producer of the excluded product cannot steal the sale simply by offering a somewhat lower price. For example, the pizza franchisees in *Queen City* or the tooth product dealers in *Dentsply* could not profit by purchasing cheaper pizza dough or tooth filling materials from a rival seller because any gains from lower prices would almost certainly not be enough to compensate them for the loss of their dealerships.[22]

THE RESULT IS THAT AN EQUALLY EFFICIENT PRODUCER OF THE EXCLUDED PRODUCT CANNOT STEAL THE SALE SIMPLY BY OFFERING A SOMEWHAT LOWER PRICE.

By contrast, the discount conditioned on exclusivity places the buyer in a much different position: when it purchases from a rival it loses the discount, but not its dealership or franchise. If Domino's merely offered its franchisees a 10 percent discount if they committed to purchasing all of their pizza dough from their franchisor, then any rival would have been able to steal the franchisees' trade simply by meeting or beating the discounted price. Because the franchise itself is not at risk, an equally efficient rival should be able to steal the sale as long as the fully discounted price is above cost. Further, because the strategy excludes only if the prices are predatory, *Brooke Group*'s recoupment requirement applies as well.

---

20 *E.g.,* United States v. Dentsply Int'l., Inc., 399 F.3d 181 (3d Cir. 2005), *cert. denied*, 126 S.Ct. 1023 (2006) [hereinafter *Dentsply*].

21 *See, e.g.,* ROGER D. BLAIR & DAVID KASERMAN, LAW AND ECONOMICS OF VERTICAL INTEGRATION AND CONTROL (1983); Paul Joskow, *Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence*, 4 J.L., ECON. & ORG. 95 (1988); Ian MacNeil, *Contracts: Adjustment of Long(Term Economic Relations Under Classical, Neoclassical, and Relational Contract Law*, 72 NW. U. L. REV. 854 (1978).

22 Queen City Pizza, Inc. v. Domino's Pizza, Inc., 124 F.3d 430 (3d Cir. 1997), *cert. denied*, 523 U.S. 1059 (1998) (refusing to condemn franchisor's requirement that franchisees purchase its own pizza dough exclusively) [hereinafter *Queen City*]; *Dentsply*, *supra* note 20 (condemning manufacturer's requirement that dealers purchase its artificial tooth material exclusively).

This analysis applies to all situations in which the discount applies to a single product, or where the discount applies to multiple products but at least one significant rival makes the same set of products. In all such cases, an equally efficient rival could steal the sale. It necessarily also applies to quantity and market-share discounts.[23] A fortiorari, a discount that requires the purchaser to take less than 100 percent of its product from the seller excludes less than a discount conditioned on exclusivity.

III. Package Discounts

A package discount is one that is aggregated across two or more distinct goods.[24] In order to have the effects associated with package discounts, the discount must not merely apply to two or more goods; it must also be aggregated across them. For example, if a seller sells widgets (A) and gadgets (B) and gives the buyer a 10 percent discount for taking at least 10,000 units of either, that is not a package discount. Failing to meet the quota on one does not impact the price of the other. A package discount would be an offer of a 10 percent discount if the buyer took 20,000 units of any combination of A and B; or alternatively, if it took at least 80 percent of its total needs of A and B from this particular seller.

Package discounts can exclude even equally efficient rivals who do not sell all of the goods in the package. For example, suppose that the dominant seller has costs of $c_A = 10$ and $c_B = 6$. It offers individual prices of $p_A = 14$ and $p_B = 8$. The seller also offers a discounted price of 19 to a buyer who takes one A and one B. Note first that the price of the package, 19, is well above the seller's costs of 16. However, while an equally efficient rival could sell B alone for 6, undercutting the seller's undiscounted price, it would not be profitable for the buyer to purchase B at 6 from the rival. The buyer would have to pay 14 for the dominant firm's product A, and the combined $p_A + p_B$ price would be 20. Indeed, the only way the rival could make the customer an attractive offer would be to sell B at a price under 5, which would be less than its costs.

While this practice could exclude particular rivals, it would be exclusionary in the antitrust sense only if no substantial rival offered the same AB package that the dominant firm did. That is, if two or more equally efficient firms offered the AB package, then a package discount to 19 would be easily met. So the strategy

23 *E.g.*, Concord Boat Co. v. Brunswick Corp., 207 F.3d 1039 (8th Cir.), *cert. denied*, 531 U.S. 979 (2000) (refusing to condemn above cost market share discounts because purchasers were free to walk away at any time and purchase from a rival).

24 *See* LePage's, Inc. v. 3M Corp., 324 F.3d 141 (3d Cir. 2003) (en banc), *cert. denied*, 540 U.S. 807 (2003) (condemning package discounts without requiring below-cost pricing) [hereinafter *LePage's*]; accord SmithKline Corp. v. Eli Lilly & Co., 575 F.2d 1056 (3d Cir.), *cert. denied*, 439 U.S. 838 (1978); cf. Virgin Atlantic Airways Ltd. v. British Airways PLC, 69 F.Supp.2d 571, 580 n. 8 (S.D.N.Y. 1999), *aff'd on other grds*, 257 F.3d 256 (2d Cir. 2001) (refusing to condemn).

requires as a minimum condition that all significant rivals offer either item A or item B, but not both.

The legal debate over package discounts has focused mainly on whether they should be analogized to predatory pricing or to tying. Defendants have generally preferred a test likening the practices to predatory pricing, which typically means proof that the price of the bundle is below the seller's marginal or average variable cost for the goods contained in the bundle, and that the defendant will be able to recoup its predation investment with higher prices once rivals are excluded from the market. By contrast, plaintiffs have likened bundled discounts to tying arrangements, which require proof that two different goods are tied together, but do not require either below-cost prices or recoupment.

The tying analogy is the better one, but a cost test is necessary to establish that the two products are indeed tied. Tying law requires that the tying and tied product be tied together, which means that the buyer has a strong incentive to take both products from the seller, thus excluding rivals in the tied product market. The most explicit tie is the contract requiring the buyer to purchase both goods, as when a fast food franchisor requires franchisees to purchase all of a certain ingredient from the franchisor or else forfeit its franchise. Some contractual tying requirements can be implicit rather than explicit, and tying must then be proven from the circumstances.[25] In other cases, the tie is "technological," as when the maker of a camera designs it in such a way that it will accept only its own film cartridges.[26] Finally, package discounts are ties when the pricing strategy makes it unprofitable for rivals to match the package discount.

Two products are said to be "tied together" by means of a package discount if one attributes the total discount to the particular good for which exclusion is claimed and the resulting price is below that good's marginal or average variable cost. In that case, a purchase of the goods separately will be more costly than purchasing them in the bundle.[27]

25 *E.g.*, Cia. Petrolera Caribe v. Avis Rental Car Corp., 735 F.2d 636, 637-638 (1st Cir. 1984) (tying or rental cars and gasoline could not be inferred merely from fact that large proportion of defendant's rental car customers also purchased gasoline from defendant); American Mfrs. Mut. Ins. Co. v. American Broadcasting-Paramount Theatres, 446 F.2d 1131 (2d Cir. 1971), *cert. denied*, 404 U.S. 1063 (1972) (simple hard bargaining where seller preferred tie and buyer acquiesced without making a counteroffer did not constitute tying); Unijax v. Champion Int'l, 683 F.2d 678, 686 (2d Cir. 1982) (mere fact that plaintiff purchased the two products together as offered did not constitute tying absent evidence that the defendant would have refused a request to sell them separately).

26 For all these variations of the requirement that two products be "tied" together, *see* 10 ANTITRUST LAW ¶¶ 1752-1758 (2d ed. 2004). *See* Berkey Photo v. Eastman Kodak Co., 603 F.2d 263, 287 (2d Cir. 1979), *cert. denied*, 444 U.S. 1093 (1980) (design required user's of new Kodak camera to use its cartridge film as well). *See also* United States v. Microsoft, 253 F.2d 34, 66-67 (D.C. Cir. 2001), *cert. denied*, 534 U.S. 952 (2001) (Microsoft's "commingling" of Windows and Internet Explorer code made it impossible for a buyer to purchase one without the other).

27 To be sure, a rival who sold both items could match the offer, but it would still be a tied offer. So while the goods are tied together, a buyer in that case could purchase from alternative sellers.

This discount attribution test establishes that two goods are tied together. Even when tying exists, however, most such arrangements are lawful because they occur in competitive markets or have perfectly innocent explanations. Packages discounts are competitively harmless if they do not exclude rivals sufficiently to facilitate the exercise of market power, if they are cost justified because joint provision is less expensive than single provision, or if joint provision improves product quality or pleases customers in other ways.[28]

Joint provision can also be a means of price discrimination. For example, suppose that the dominant seller sells A and B with costs $c_A = 6$ and $c_B = 8$. Buyers X and Y both value A and B, but by different amounts:

X's reservation prices are $p_A = 7$; $p_B = 12$

Y's reservation prices are $p_A = 10$; $p_B = 9$

If the seller sold the goods separately it could set prices of $p_A = 7$ and $p_B = 9$, selling two of each A and B and earning total profits of 4. Alternatively, it could sell 1 A at $p_A = 10$ and 1 B at $p_B = 12$, earning total profits of 8. The seller's best alternative would be to sell one of each at the higher prices. However, if the seller charged $p_{AB} = 19$ for the package, both buyers would purchase both products, and the seller would earn profits of 10.

While this illustration can be written an infinite number of ways, it shows that bundling can be output-increasing even though it results in higher economic profits to the seller. In the above illustration, the seller who is prohibited from offering a bundled discount would sell individually at prices of $p_A = 10$ and $p_B = 12$, and output would be half as high as with a bundled discount price of $p_{AB} = 19$.

No good rationale exists for condemning output increasing practices under the antitrust laws. More significantly, the profitability of bundling used to achieve price discrimination does not depend on the exclusion of any rival. The only objection to the practice under these circumstances is that it extracts more consumers' surplus than single product pricing would, or perhaps that the transaction costs of a price discrimination scheme exceed any gains that price discrimination might produce. But these are certainly not warrants for condemning a practice as exclusionary under Section 2 of the Sherman Act.

The case law on bundled discounts has just begun to scratch the surface of these issues. Although the opinion is unclear, the *LePage's* case condemned bundled discounts without a showing that an equally efficient rival could not match the discounts. Several district court decisions have assessed the basic require-

28 *See, e.g.*, David S. Evans & Michael Salinger, *Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law*, 22 YALE J. ON REG. 37 (2005). They can also be efficient under the same general conditions that tying or exclusive dealing are efficient. On tying, *see* 9 ANTITRUST LAW ¶¶ 1711-1718 (2d ed. 2004); on exclusive dealing, *see* 11 ANTITRUST LAW ¶¶ 1810-1814 (2d ed. 2005).

ment that when the discount is fully attributed to the exclusion product the price of that product must fall below cost.[29] In general, however, the decisions have not reached beyond this to analyze why even bundled discounts that meet this test might be beneficial or competitively benign.

A sensible legal test for unlawful package discounts would proceed in this way. If the defendant had one or more rivals that produced the same goods as are included in the package discount, then the package would be unlawful only if the package as a whole was sold at a price less than the relevant measure of cost—that is, the general test for predatory pricing would apply. By contrast, if the defendant was the only firm that offered the discounted package, the analysis would proceed in two steps. First, a price/cost analysis would have to be used to establish that the products in the bundle are actually tied together. Tying would occur if, when all discounts are attributed to the product upon which exclusion is claimed, the price of that product falls below the relevant measure of cost. If the overall price of the bundle exceeds cost, however, then all the defenses normally applied in tying arrangement cases would apply here as well. Bundling is generally pro-competitive if it reduces transaction costs, results in improved products or services, or enables quality control.[30]

BUNDLING IS GENERALLY PRO-COMPETITIVE IF IT REDUCES TRANSACTION COSTS, RESULTS IN IMPROVED PRODUCTS OR SERVICES, OR ENABLES QUALITY CONTROL.

Even this test would have to be qualified in a market in which some rivals produced only the A product and others produced only the B product. In that case, a pair of rivals could join together and match the dominant firm's package discount.[31] This could also happen if a broker or other intermediary assembled goods from numerous sellers and was able to offer a package on terms equivalent to those being offered by the defendant.

29 *See* Information Resources, Inc. (IRI) v. Dun & Bradstreet Corp., 359 F.Supp.2d 307, 307-308 (S.D.N.Y. 2004), which concluded:

> When price discounts in one market are bundled with the price charged in a second market, the discounts must be applied to the price in the second market in determining whether that price is below that product's average variable cost.

At this writing, this decision is on appeal to the Second Circuit. Accord Virgin Atlantic Airways Ltd. v. British Airways PLC, 69 F.Supp.2d 571, 580 n. 8 (S.D.N.Y. 1999), *aff'd on other grds*, 257 F.3d 256 (2d Cir. 2001).

30 These and other pro-competitive rationales for tying are discussed in 9 ANTITRUST LAW ¶¶ 1711-1718 (2d ed. 2004).

31 *See* Thomas A. Lambert, *Evaluating Bundled Discounts*, 89 MINN. L. REV. 1688, 1742, 1746-1748 (2005).

Indeed, in a well-functioning market containing equally efficient rivals producing each of the two goods, explicit coordination between makers of product A and product B is unnecessary. The rival seller of product A would realize that its viability depends on its charging a price for A that is low enough so that B can also be viable, and vice versa. As a result, each would cut its price so as to accommodate the other. To illustrate, suppose that the dominant firm produces good A, with costs of $c_A = 10$ and a price of $p_A = 13$; and good B, with costs of $c_B = 5$ and a price of $p_B = 8$. It offers a package discount price of $p_{AB} = 17$. An equally efficient rival in product A cannot match the discount, because its cost price of $c_A = 10$ and the dominant firm's price of $p_B = 8$ for B would be too high. An equally efficient rival in B cannot match the discount because its cost of $c_B = 5$ for B and A's price of $p_A = 13$ would also be too high. However, it would be feasible for the buyer to purchase product A from a rival at a price of $p_A = 10$, and product B from a different rival at a price of $p_B = 8$. It would be in both of these sellers' best interest to cut their prices sufficiently to enable the buyer to do this.

IV. Predatory Purchasing

Exclusionary purchasing occurs when a dominant firm pursues a strategy of buying up so much of some input that rival firms cannot obtain adequate access. The classic form of such strategies was the output contract. For example, American Can contracted with all of the then existing makers of patent can making machinery to sell their total output of machinery to American Can.[32] Assuming that such agreements were anticompetitive, the machine makers would agree to such contracts only if they were compensated, perhaps with higher prices.[33] Of course, most output contracts have competitively benign explanations. Most occur in competitive markets and are mechanisms by which firms assure themselves of a reliable source of supply.[34]

Two forms of predatory purchasing that the case law has analogized to predatory pricing are slotting fees and "predatory spending," which refers to a group of strategies of overinvesting in productive capacity, or else buying up scarce inputs at a high price in order to deny market access to rivals.

32 United States v. American Can Co., 230 F. 859 (D. Md. 1916), *appeal dismissed*, 256 U.S. 706 (1921). *See* 11 ANTITRUST LAW ¶ 1801a (2d ed. 2005).

33 *See* Eric B.. Rasmusen, J. Mark Ramseyer, & John S. Wiley, Jr., *Naked Exclusion*, 81 AM. ECON. REV. 1137 (1991); Thomas G. Krattenmaker & Steven Salop, *Anticompetitive Exclusion: Raising Rivals' Costs To Achieve Power over Price*, 96 YALE L.J. 209, 236-238 (1986).

34 *See* 11 ANTITRUST LAW ¶¶ 1810-1814 (2d ed. 2005).

A. SLOTTING FEES

A "slotting fee" is a payment made by a manufacturer to a dealer, typically a retailer, in exchange for guaranteed display space or some other preferential treatment of the manufacturer's product. For example, a seller of spices might give a retailer $1,000 up front in exchange for a year's access to five linear feet of retail shelf space. Given that shelf space in desirable stores is in fiercely short supply, such fees can be exclusionary in the sense that the retailer stocks the payor's product to the exclusion of someone else's. Indeed, if there were plenty of empty shelf space, slotting fees would not exist.

The main function of slotting fees is to transfer a portion of the risk of poor sales from the retailer to the supplier. The slotting fee, which is fixed, operates as a discount whose size varies inversely with the volume of goods that the retailer sells in that space. For example, if spices sell for $1.39 per bottle and the retailer sells only 1,000 bottles per year in the allocated space, the $1,000 slotting fee operates as a prohibitively high $1.00 per bottle discount on the price. However, if the retailer sells 100,000 bottles in that space, the discount is only $.01 per bottle, but the retailer is more than happy because of the high sales volume. Thus, the willingness to pay slotting fees operates as a signal to the retailer that the supplier has confidence in its product. Alternatively, the fee compensates the retailer if the product ends up not doing very well. In sum, slotting fees have strong pro-competitive benefits and the economic case against them is very weak.[35]

Slotting fees cannot exclude an equally efficient rival unless they are so high that they drive the product price below cost.[36] The relevant measure is reasonably anticipated cost at the time the slotting fee is negotiated. Firms should not offer such efficiency enhancing arrangements at their peril if it later turns out that sales did not materialize as the manufacturer hoped. Indeed, the purpose of the slotting fee is to shift the risk of poor sales to the manufacture, and the fact that the fee is required is good evidence that there is, in fact, some risk. The real question is whether the manufacturer had a reasonable, objectively measurable expectation ex ante that the product would produce sufficient sales so that the price net of the slotting fee would be profitable to the manufacturer.

35 *See* 11 ANTITRUST LAW ¶ 1807c (2d ed. 2005); *see also* REPORT ON THE FTC WORKSHOP ON SLOTTING ALLOWANCES AND OTHER MARKETING PRACTICES IN THE GROCERY INDUSTRY (U.S. Federal Trade Commission, Washington, DC, 2001); Mary W. Sullivan, *Slotting Allowances and the Market for New Products*, 40 J.L. & ECON. 461 (1997).

36 *See* El Aguila Food Products, Inc. v. Gruma Corp., 301 F.Supp.2d 612, (S.D.Tex. 2003), *aff'd mem.*, 131 Fed.Appx. 450, 2005-1 Trade Cas. ¶ 74788 (5th Cir. May 17, 2005) (refusing to condemn slotting fees as monopolistic when prices were above any relevant measure of cost).

B. PREDATORY SPENDING

Predatory spending is the inverse of predatory pricing. The defendant monopsonist, who may or may not have market power in the output market, pays a high price for some scarce input with the result that rivals are unable to obtain it and are driven from business. Anticompetitive predatory spending has welfare effects similar to predatory pricing. The risks of overdeterrence and false positives are equivalent to those in predatory pricing cases. Further, the claims are even harder to evaluate, magnifying the possibility of error.

The decisions can be roughly grouped into two types—although there is considerable overlap between them. One type is best termed predatory investment, and refers to situations in which the defendant allegedly overbuilds or overinvests in facilities in order to deny market access to rivals. These cases themselves have come in numerous varieties, including claims that the defendant invested in a larger plant than it needed,[37] that it built excessive retail facilities, setting them up in such a way as to deny rivals adequate market access,[38] or similarly, that an airline responded to a rival's entry by flooding the market with additional aircraft.[39] The second type of case, which is much more analogous to conventional predatory pricing, claims that the defendant engaged in overbuying of some variable cost input into its production process.[40]

The predatory investment cases are not as easily classified or characterized as the traditional predatory pricing cases. In some, such as *DuPont*, the excess investment in a production facility may have been a form of strategic entry deter-

37 DuPont (Titanium-Dioxide), 96 F.T.C. 653 (1980) [hereinafter *DuPont*]. *See* 3 ANTITRUST LAW ¶ 737d (2d ed. 2002).

38 Photovest Corp. v. Fotomat Corp., 606 F.2d 704 (7th Cir. 1979), *cert. denied*, 445 U.S. 917 (1980) (defendant overbuilt retail outlets in excess of the number its own studies indicated were necessary).

39 *Spirit Airlines*, *supra* note 2 (separate portion of opinion denying summary judgment on claim of predatory transfer of aircraft into the plaintiff's routes, without requiring prices below cost).

40 *E.g.*, Confederated Tribes of Siletz Indians of Oregon v. Weyerhaeuser Co., 411 F.3d 1030 (9th Cir. 2005), *pet. for cert. filed*, ___ S.Ct. ___ (2005) [hereinafter *Weyerhaeuser*]; Reid Bros. Logging Co. v. Ketchikan Pulp Co., 699 F.2d 1292, 1298 n. 5 (9th Cir. 1983). *See also* Am. Tobacco Co. v. United States, 328 U.S. 781, 801-04 (1946) (defendants conspired to buy up cheaper tobacco at high prices in order to deny access to rivals who were making lower cost cigarettes); United States v. Aluminum Co. (Alcoa), 148 F.2d 416, 432-433 (2d Cir. 1945) (referring to claims that Alcoa had bought up bauxite and electric power in order to deny access to rivals). Cf. Syufy Enters. v. American Multicinema, 793 F.2d 990 (9th Cir. 1986), *cert. denied*, 479 U.S. 1031 (1987) (defendant purchased excessive exclusive licenses for exhibiting films, thus denying access to rival exhibitors; affirming judgment for plaintiff); contrast Houser v. Fox Theatres Management Corp., 845 F.2d 1225 (3d Cir. 1988) (rejecting claim that defendant exhibitor overbought exclusive bookings on films in order to deny access to rivals); Potters Med. Center v. City Hosp. Assn., 800 F.2d 568 (6th Cir. 1986) (denying summary judgment on claim that defendant used salary guarantees and other perks to induce physicians to accept exclusive privileges at its hospital).

rence. By building a very large plant with well publicized excess capacity, the dominant firm could threaten new entrants with an immediate output increase and price reduction, thus inducing them not to enter in the first place. As a dominant firm, it could still set price significantly above its costs, including the costs of carrying the additional capacity.

As the outcome exonerating the defendant in the *DuPont* case suggests, taking long-run concerns into account in predatory investment cases is just as difficult as it is in orthodox predatory pricing cases. With respect to sale prices, long-run concerns arise mainly with respect to claims that prices above average total cost are predatory. Even though such prices produce short-run profits, the claim is that an even higher price would be more profitable over a longer run—or more specifically, that the price is profitable mainly as an entry deterrence device. In *Brooke Group*, the U.S. Supreme Court categorically rejected such claims. As a result, prices below cost are an essential element of a predatory pricing case.

Long-run concerns become relevant in predatory spending situations when it is claimed, for example, that the firm invested in a much larger plant than it needed in order to deter entry; that it intentionally targeted markets occupied by weak rivals in deciding where to deploy assets, and the like. The prices are presumably sustainable and the defendant is earning a profit; however, it has overbuilt its capacity in some way that excludes rivals and thus permits a longer stream of monopoly profits sufficient to offset the additional development costs.

Long-run concerns can blend into short-run when durable assets are readily transferable. In the *Spirit Airlines* case, the U.S. Court of Appeals for the Sixth Circuit accepted the plaintiff's argument that the defendant, Northwest Airlines, shifted aircraft into the plaintiff's markets, dropping the price precipitously, only to shift the aircraft back once the plaintiff had been forced out.[41] The court permitted this claim to go to trial even when there was no showing that the result of the shift was to drive the defendant's selling prices below its costs. Assuming the fact finder properly considered the opportunity cost of the lost revenue on the routes from which the aircraft were transferred, the claim must have been that the defendant shifted aircraft from more profitable to less profitable routes and that this strategy made sense only because of its value in knocking a rival out of the market so that the defendant could thereafter raise price in the targeted routes and recover its investment.

Shifting of aircraft is a more aggressive and more costly act than cutting a price. As a result, a court might be more comfortable condemning a two-way shift ("in" when in response to new entry and "out" after the rival has been excluded) without a showing of prices below cost. But the shift would still have to be costly in the short run. Otherwise, it would be perfectly rational conduct without regard to exclusion of rivals. Thus, proof of recoupment seems essential, unless perfectly

41 *See Spirit Airlines*, *supra* note 2.

appropriate competitive behavior is to be condemned. While the Sixth Circuit did not require proof of below-cost prices, it did require proof of recoupment.[42]

The purest short-run predatory spending strategy involves the dominant firm that purchases variable cost inputs at a predatorily high price, thus making them unavailable to weaker rivals. The U.S. Court of Appeals for the Ninth Circuit's recent *Weyerhaeuser* decision involved a defendant who was dominant in the purchasing market for Alder logs, but which sold lumber made from the logs in a competitive market.[43] The logs themselves were the principal input into the lumber. Approximately 75 percent of the finished lumber's cost consisted of the price paid for the logs. The defendant allegedly engaged in "overbuying," that is purchasing all the logs that it could at unreasonably high prices with the result that rival sawmills were unable to purchase enough logs or to make enough margin on the logs that they processed, thus driving them out of business.

THE JURY INSTRUCTION THAT THE NINTH CIRCUIT APPROVED IS AN ANTITRUST DISASTER OF ENORMOUS PROPORTIONS.

The Ninth Circuit upheld a judgment for the plaintiff based on an instruction entitling the jury to find a Sherman Section 2 violation if the defendant paid "too much," or "more than necessary," for the logs. The court held that such a standard was sufficient, and it did not require the plaintiff to show either that the purchase price was so high as to drive the defendant's resale price for finished lumber below its cost, or that the defendant would be able to recoup its investment in high priced logs by paying lower prices after rival sawmills were excluded from the market.

Three observations seem relevant. First, the jury instruction that the Ninth Circuit approved is an antitrust disaster of enormous proportions. Short-run supply bottlenecks are relatively common, and price is the principal rationing device for scarce inputs. Large buyers subject to Ninth Circuit law now operate under the threat that if they bid too aggressively for some scarce input a jury will find that they paid "more than necessary" and subject them to treble damage liability. There is no obvious reason for thinking this will be a rare occurrence. Some kind of standard with more substance is essential.

---

42 *Id.*, at para. 32:

> . . . [E]ven if the jury were to find that Northwest's prices exceeded an appropriate measure of average variable costs, the jury must also consider the market structure in this controversy to determine if Northwest's deep price discounts in response to Spirit's entry and the accompanying expansion of its capacity on these routes injured competition by causing Spirit's departure from this market and allowing Northwest to recoup its losses and to enjoy monopoly power as a result.

43 *See Weyerhaeuser*, *supra* note 40.

Second, the Ninth Circuit's refusal to require proof that the defendant's sales were below a relevant measure of cost is incorrect in a market where costs are easily defined and make up a significant portion of the purchase price—in this particular case, some 75 percent of the cost of the finished lumber. In such a situation, it should be quite easy to conclude that an input purchase price is not "too high" unless a firm is unable to make a profit on its sales. That seems doubly true in a case such as *Weyerhaeuser*, where the defendant sold lumber in a competitive market and had no pricing discretion. In such circumstances, it is economically impossible to say that a defendant is paying too much for an input if it is earning a competitive return on what it sells.

On administrative grounds, a price/cost test is more difficult to defend if the input in question constitutes only a small percentage of the cost of the finished product. For example, suppose that hardwood saw blades were in short supply and Weyerhaeuser acquired them by bidding up the price. Suppose that a saw blade is a variable cost item because it wears out and its cost amortizes out at less than $^1/_2$ percent of the total cost of the finished lumber. Even if the defendant paid double the market price for saw blades, the difference is likely to be within its margins. It would be almost impossible to show that overpaying for saw blades drove the defendant's prices below its costs. In such cases, courts might need to look for other hard evidence of exclusionary behavior. For example, the defendant might have purchased saw blades and stockpiled them for very long periods or even destroyed them, simply to deny access to rival sawmills. However, even here the courts must be careful. For example, stockpiling of inventories in times of anticipated shortages is perfectly pro-competitive behavior. A firm that has a reasonable expectation at the time of purchase that it actually will use an input in its own production should never be condemned for behaving predatorily. In any event, the fact findings here were that Weyerhaeuser was reselling the finished lumber in a competitive market.[44] In that case, it could have sold all it wanted at the competitive price. For the same reason, such a firm would have no incentive to overbuy and destroy the excess—in a competitive resale market there would be no excess.

Third, proof of recoupment seems essential in all cases of predatory spending. Whether or not the defendant's costs are pushed higher than its prices, an anticompetitive strategy of overbuying will not be profitable unless its payoff is greater than the investment. In the great majority of cases what appears to be overbuying will be nothing more than hedging against an uncertain future. For example, the computer manufacturer that stockpiles RAM chips in contemplation of a possible future shortage is simply engaging in self protection.

44 The most likely explanation for significant power in the buying market and lack of power in the selling market in this case was that the economies of buying and shipping raw logs limited the geographic market on the purchasing side to a fairly small range. All of the sawmills were located close to the producing forests. However, the finished lumber was sold in a market that was nationwide or even larger. *See, e.g.*, Mandeville Island Farms v. Am. Crystal Sugar Co., 334 U.S. 219 (1948) (seller purchased sugar beets in small geographic area but sold refined sugar in national market); 12 Herbert Hovenkamp, Antitrust Law ¶ 2011b (2d ed. 2006

V. Conclusion

While success in conventional predatory pricing cases has been elusive for plaintiffs since the U.S. Supreme Court's *Brooke Group* decision in 1993, they have had better luck with variant predatory pricing or spending practices. It is hard to avoid the conclusion, however, that the relatively greater success results from the fact that these practices are not very well understood. When judges do not understand practices very well, they tend to give them to juries, and juries often find for plaintiffs, particularly if the defendant's intentions seemed to be anticompetitive. As a result, one can predict that as variant predatory pricing practices are understood more fully, plaintiffs' success rates will decline in these areas as well. This is not to say that the set of legitimate anticompetitive pricing claims is empty, but that the existence of predation that is within the competence of courts to evaluate remains very rare.

Volume 2 | Number 1 | Spring 2006

Competition Policy International

Article 82: A Commentary on DG Competition's Discussion Paper

Bill Allan

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Article 82: A Commentary on DG Competition's Discussion Paper

Bill Allan

DG Competition's discussion paper is a welcome commitment to a consumer welfare standard implemented through an effects-based control of exclusionary abuses. As such, it appears to signal a departure from the form-based approach articulated most strongly in *Michelin II*. However, its full significance is limited by the enunciation of a precautionary principle under which abuse is framed to capture any conduct likely to limit entry or expansion and justification is limited to the narrowest plausible extent. While that approach reflects a concern to prevent the erection of artificial entry barriers, it results in rules that undervalue existing competition. That risk is compounded by a narrow approach to market definition and dominance. These problems will only be avoided if the European Commission fully embraces a standard based on a determination that the disputed conduct substantially lessens effective competition in a way that can effectively be remedied by intervention under Article 82 of the EC Treaty.

The author is a lawyer with Linklaters (London) and an Affiliated Lecturer in the Faculty of Law, Cambridge University. He is indebted to his colleagues Christian Ahlborn and David Bailey for making numerous positive suggestions and for saving him from error. Any errors that remain are, of course, his own.

I. Introduction

Among the numerous opportunities for Christmas displacement activity thoughtfully provided by the competition courts and agencies, DG Competition's (DG COMP) discussion paper (Discussion Paper) was among the most widely anticipated.[1] Its appearance marks an important point in a debate that has intensified since the judgment of the European Court of First Instance (CFI) in *Michelin II*, where the Court said that:

> "The [anti-competitive] effect referred to in the case law . . . does not necessarily relate to the actual effect of the abusive conduct complained of. For the purposes of establishing an infringement of Article 82 EC, it is sufficient to show that the abusive conduct of the undertaking in a dominant position tends to restrict competition or, in other words, that the conduct is capable of having that effect.
>
> . . .
>
> It follows that, for the purpose of applying Article 82 EC, establishing the anti-competitive object and the anti-competitive effect are one and the same thing . . . If it is shown that the object pursued by the conduct of an undertaking in a dominant position is to limit competition, that conduct will also be liable to have such an effect."[2]

That approach evoked a strong reaction from those for whom it gave primacy to form over substance and, in so doing, produced results that lacked economic logic.[3]

The Discussion Paper indicates DG COMP's proposed approach to resolve that conflict. On any view, the scale of the task is daunting, dwarfing the (already accomplished) updating of Article 81 of the EC Treaty and the EC Merger Regulation and reflecting the profundity of the conflicts that the control of market power evokes. Moreover, the Discussion Paper is only a first step in the process. Its self-imposed limitation to exclusionary abuses means that it cannot

1 European Commission, DG Competition Discussion Paper on the Application of Article 82 of the Treaty to Exclusionary Abuses (Dec. 2005), *available at* http://europa.eu.int/comm/competition/antitrust/others/discpaper2005.pdf [hereinafter Discussion Paper].

2 Case T-203/01, Manufacture française des pneumatiques Michelin v. Commission, 2003 E.C.R. II-4071 (CFI) [hereinafter *Michelin II*], at paras. 239 and 241.

3 That concern was well articulated in J. Vickers, *Abuse of market power*, European Association for Research in Industrial Economics, Berlin, Sep. 3, 2004, *available at* http://www.oft.gov.uk/NR/rdonlyres/948B9FAF-B83C-49F5-B0FA-B25214DE6199/0/spe0304.pdf.

even suggest a complete resolution of conflicts that derive, to a substantial extent, from the differing perspectives of exclusionary and exploitative abuses. The development of a coherent policy that embraces and reconciles both categories of abuse is effectively deferred to the next phase in DG COMP's program.[4] Ultimately, the effect of this and any subsequent work undertaken by DG COMP depends on the extent to which it is embraced by the EC Courts.[5]

This paper is organized in the following sections. Section II addresses issues of general principle and the general analytical framework. Sections III, IV, V, and VI consider the Discussion Paper's treatment of specific abuses (predation, rebates, tying, and refusal to supply).[6] Finally, Section VII contains some concluding remarks.

II. General Principles

A. POLICY OBJECTIVES IN THE ENFORCEMENT OF ARTICLE 82 OF THE EC TREATY

1. Consumer Welfare as the Primary Goal

The Discussion Paper proclaims its central orientation in the introduction: "With regard to exclusionary abuses the objective of Article 82 is the protection of competition on the market as a means of enhancing consumer welfare and of ensuring an efficient allocation of resources."[7] The clarity of that statement is qualified by the introductory reference to exclusionary abuses.[8] While that may be no more than a precautionary qualification to avoid pre-empting the next

4 *See Commission discussion paper on abuse of dominance - frequently asked questions*, *at* http://www.europa.eu.int/rapid/pressReleasesAction.do?reference=MEMO/05/486 (last visited Feb. 7, 2006), *cited in* Discussion Paper, *supra* note 1, at 1 n.1. In defining the scope of the Discussion Paper, it should also be noted that, while it discusses issues presented by collective dominance, the attention that they receive is slight by comparison with the attention devoted to the central topic of single firm dominance and largely consists in a restatement of the principles enunciated by the EC Courts. Accordingly, that topic is not discussed in this paper.

5 *See* Case C-95/04 P, British Airways v. Commission, AG Opinion (Feb. 23, 2006) [hereinafter *Virgin/British Airways (ECJ)*], at para. 28.

6 The Discussion Paper also includes a section on aftermarkets (*see* Discussion Paper, *supra* note 1, at paras. 243-265). The discussion focuses almost entirely on the question of whether the supplier of the primary product holds a dominant position in relation to its secondary products. Where dominance is established, DG COMP says that it will presume that the supplier abuses that dominant position if it reserves the secondary market to itself. For more detailed analysis, it simply cross-refers to the sections on tying and refusal to supply.

7 Discussion Paper, *supra* note 1, at para. 5.

8 *See also id.* at paras. 54 and 56 (which include the same qualification).

phase in DG COMP's program, equally it may signal the existence of a different objective where exploitative practices are concerned.

From its inception, EC competition law has pursued a diversity of objectives that cannot always be reconciled. Those tensions crystallize when notions of fairness embodied in the protection against discrimination collide with notions of maximizing consumer welfare such as those articulated in the Discussion Paper.[9] The law cannot equivocate on this issue. When an authority or a court decides whether to apply Article 82, either it makes its decision by reference to rules designed to maximize consumer welfare or it does not. If it does not, because it applies rules designed to achieve another objective, consumer welfare is necessarily diminished.[10]

2. DG COMP's Application of the Consumer Welfare Standard to Exclusionary Abuses

As importantly, DG COMP commits itself to a methodology based on the disputed conduct's likely effect on the market where:

> "... The conduct in question must in the first place have the capability, by its nature, to foreclose competitors from the market. To establish such capability it is in general sufficient to investigate the form and nature of the conduct in question. *It secondly implies that, in the specific market context, a likely market-distorting foreclosure effect must be established.*" (emphasis added)[11]

Although not said in so many words, this should mark the welcome repudiation of the form-based philosophy articulated by the CFI in *Michelin II*.[12] The

9 The (cautiously contemplated) meeting competition defense provides an instance of that tension. It requires that the dominant firm's response be the minimum required to protect its commercial interests yet that may necessitate otherwise unjustified discrimination between its customers. DG COMP does not articulate in this paper how it proposes to resolve that tension. More broadly, EC competition law has yet to reach the conclusion, enunciated by the U.S. Supreme Court in Brooke Group, Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993), that prohibitions on discriminatory conduct only offend the antitrust laws where they interfere with competition.

10 That does not exclude other objectives (such as liberalization or single market integration) where they are wholly consistent with the consumer welfare goal.

11 Discussion Paper, *supra* note 1, at para. 58.

12 Note, however, that DG COMP enters a caveat in respect of conduct that is "clearly not competition on the merits" (*id.* at para. 60). Such conduct is presumed to be abusive subject to rebuttal evidence that the conduct has no exclusionary effect or meets the objective justification standard. The low threshold for exclusionary effect and the high threshold for justification combine to make the prospect of successful rebuttal remote. In any event, the necessity for any such qualification is questionable. If the conduct is generically as devoid of redeeming features as DG COMP suggests, the second

opinion of Advocate General Kokott in *Virgin/British Airways (ECJ)* is consistent with that position.[13]

That said, even if an effects-based approach does become firmly entrenched, its significance depends critically on the way in which it is applied. It is notable, for example, that the CFI's judgment in *Virgin/British Airways (CFI)*, endorsed by the Advocate General on appeal, reached a finding of abuse without having to rely on the full force of the *Michelin II* formula. It was sufficient for that Court to find an anticompetitive effect on the basis of a factual assumption that, absent the incentive schemes employed by British Airways (BA), rival airlines would have expanded more vigorously than they did.[14] As stated, that factual assumption had the legal effect of a presumption that is not substantially different from the *Michelin II* formula.

That leads directly to the central question of whether the scope and application of the exclusionary abuses will be determined by reference to the disputed conduct's effect on entry and expansion by rivals alone or on a broader basis in which those factors are treated as part of an assessment of the disputed conduct's effect on the intensity of competition. Although the Discussion Paper includes elements of both approaches, its dominant philosophy may be described as a precautionary principle under which any threat to the long-term competitive structure of the market is sufficient to justify intervention.[15] That underpins an analytical framework comprising a broadly defined concept of foreclosure and a nar-

footnote 12 cont'd

(market-specific) limb of the inquiry should not be unduly burdensome. Moreover, if these cases are as infrequent as DG COMP implies, then the need to adopt a per se standard to avoid substantial administrative burdens seems slight. At a minimum, it is important that this exception be confined to the margins that its location in the Discussion Paper indicates. To ensure that that happens, it would be useful if the final version of the Discussion Paper were to state expressly that this caveat does not apply to the principal abuses covered by the paper. Given the recognition that each of these abuses is capable of having beneficial effects, they would seem inappropriate candidates for this treatment.

13 *See Virgin/British Airways (ECJ)*, AG Opinion, *supra* note 5, at para. 45.

14 Case T-219/99, British Airways v. Commission, 2003 E.C.R. II-5917 (CFI) [hereinafter *Virgin/British Airways (CFI)*], at para. 298. In *Virgin/British Airways (CFI)*, para. 239 of *Michelin II* was repeated, but para. 241 was not. *See also Virgin/British Airways (ECJ)*, AG Opinion, *supra* note 5, at para. 83.

15 The Advocate General's opinion in *Virgin/British Airways (ECJ)* expressly endorses that approach. See, e.g., *Virgin/British Airways (ECJ)*, AG Opinion, *supra* note 5, at para. 68 (". . . Article 82, like the other competition rues of the Treaty, is not designed only or primarily to protect the immediate interests of individual competitors or consumers, but to protect the *structure of the market* and thus *competition as such* (*as an institution*) . . . In this way, consumers are also indirectly protected. Because where competition is damaged, disadvantages for consumers are also to be feared" (emphasis in original, footnotes omitted)).

rowly circumscribed set of justifications in which primacy is given to ensuring that entry and expansion is possible.[16]

THE DISCUSSION PAPER'S DOMINANT PHILOSOPHY MAY BE DESCRIBED AS A PRECAUTIONARY PRINCIPLE UNDER WHICH ANY THREAT TO THE LONG-TERM COMPETITIVE STRUCTURE OF THE MARKET IS SUFFICIENT TO JUSTIFY INTERVENTION. THAT UNDERPINS AN ANALYTICAL FRAMEWORK COMPRISING A BROADLY DEFINED CONCEPT OF FORECLOSURE AND A NARROWLY CIRCUMSCRIBED SET OF JUSTIFICATIONS IN WHICH PRIMACY IS GIVEN TO ENSURING THAT ENTRY AND EXPANSION IS POSSIBLE.

It appears that, for DG COMP, its concept of foreclosure is saved from being overly inclusive by the (frequently repeated) proposition that, at least when considering pricing abuses, it is only the foreclosure of "as efficient" competitors that should engage Article 82.[17] Even if that is accepted at face value, it does not answer the central question of whether Article 82's guiding philosophy consists in the protection of competition as a structural phenomenon or a process of rivalry. It is consistent with either approach.

In any event, the significance of the concept is qualified (possibly substantially) by the pregnant note that "it may sometimes be necessary in the consumers' interest to also protect competitors that are not (yet) as efficient as the dominant firm."[18] It is far from clear how DG COMP intends to apply this qualification.[19] There is a significant risk that a policy that measures competitive health by long-term structural factors will deny

16 See, for example, the discussion of conditional rebates, where DG COMP says that: "In its assessment the Commission will in particular be attentive that the rebate system does not foreclose potential competitors" (Discussion Paper, *supra* note 1, at para. 157).

17 *Id.* at para. 63. It is this notion that enables DG COMP to say that Article 82 is concerned with the protection of competition not competitors. The philosophical difficulty that that disjunction can present is neatly illustrated by the interim measures rulings in *IMS Health* where the President of the CFI asserted that Article 82 was concerned with safeguarding the interests of consumers rather than protecting the position of individual competitors. The President of the ECJ responded that that approach could not be accepted without reservation "in so far as it could be understood as excluding protection of the interests of competing undertakings from the aim pursued by Article 82 EC, even though such interests cannot be separated from the maintenance of an effective competition structure." Case C-481/01, NDC Health v. Commission, 2002 E.C.R. I-3401 (ECJ), at para. 84 on appeal from Case T-184/01, IMS Health v. Commission 2001 E.C.R. II-2349 (CFI) [hereinafter *IMS Health*]. To the same effect, see the extract from the Advocate General's opinion in *Virgin/British Airways (ECJ)* quoted in *supra* note 15.

18 Discussion Paper, *supra* note 1, at para. 67.

19 Although DG COMP does not specify when that may be necessary, it appears that at least one set of cases will be those in which there are economies of scale and scope, learning curve effects, or first mover advantages that a rival could not match even if it achieved the same scale of output as the dominant firm. The framing of the qualification is curious. It implies that the assessment should be made on the hypothesis that the rival achieves comparable scale yet refers to advantages (including specifically economies of scale) that should have been eroded by the time that comparable scale is

consumers the immediate benefit of lower prices based on a dominant firm's superior economies of scale and scope or impose on consumers the costs of supporting inefficient entry.

Indeed, it is striking that there is no express reference to the merits of productive efficiency in DG COMP's recital of virtuous competitive objectives.[20] Especially where economies of scale or scope are substantial, there is an unavoidable tension between optimizing the gains from productive efficiency and facilitating entry and expansion. The Discussion Paper clearly resolves that tension in favor of entry and expansion, arguing that, in the long run, consumers must benefit more from the maintenance of a competitive structure.[21] The implications of that approach appear most starkly when DG COMP denies a dominant firm the right to justify conduct that has the effect of eliminating competition even if that conduct is necessary to achieve consumer benefits that would not otherwise be realized.

3. An Alternative Approach

DG COMP's vision for the future of Article 82 is based on the laudable policy of promoting consumer welfare through an effects-based analytical model. To accomplish that vision, the rules that implement it must be consistent with it. For the reasons explained above, rules based on a precautionary principle do not satisfy that requirement. To the contrary, there is a substantial risk that the precautionary principle will lead the European Commission to place undue weight on the assumed gains from long-term improvements in structure at the expense of the arguably more tangible gains from short-term dynamism.

Any rule should seek to avoid a result that encourages entry or expansion at the expense of weakening the intensity of competition among existing rivals.[22] To

footnote 19 cont'd

achieved. A more comprehensible hypothesis would make the assumption that the rival can match the dominant firm's efficiency once it reaches comparable scale, but is precluded from doing so by its present lack of scale. If, to the contrary, DG COMP means literally what it says, that only reinforces the concerns expressed in the text.

20 Discussion Paper, *supra* note 1, at para. 5. This approach contrasts sharply with the philosophy that: "Antitrust aims at preserving competition as an instrument for creating economic efficiency" (Frank H. Easterbrook, *The Limits of Antitrust,* 63 Texas L. Rev. (1984), *reprinted in* 1 Competition Pol'y Int'l 179, 190 (Spring 2005). However, it is consistent with EC competition law's focus on consumer welfare in preference to total welfare under which little or no value appears to be assigned to gains in productive efficiency that are retained as part of the producer surplus.

21 Discussion Paper, *supra* note 1, at para. 91.

22 Admittedly, DG COMP refers at various points to the need to take account of rivals' ability to counteract the dominant firm's conduct but, when viewed in the context of the Discussion Paper as a whole, those references are cursory and undeveloped. They do little to counterbalance the powerful thrust in the opposite direction.

avoid that result while achieving the goals of Article 82, the Commission should base its assessment on whether the disputed conduct gives rise to a substantial lessening of competition that can be effectively avoided by the proposed remedy.

This is particularly important for markets where there is already active competition, even if it is impaired by the presence of the dominant firm. In such cases, the assessment of foreclosure should concentrate on actual exclusion and deterred investment to expand. Entry and expansion are present realities; the fact that appreciable competition already exists should make a competition authority question whether entry is in fact as difficult as may be claimed.[23] More importantly, the Commission should take full account of rational competitive behavior among established competitors. If it did so, it would not place the emphasis that it does, at several points, on a criterion related to the dominant firm's coverage of total costs. Rational pricing decisions for competitors of all sizes should disregard sunk costs with the result that there are numerous instances in which prices would be set at a level below total costs. The decision-making process would start from the proposition that such pricing works to the benefit of consumers, unless the contrary is established, rather than presuming that harm will result.

This approach would also remove at least some of the difficulties presented by DG COMP's strict approach to the available defenses to an alleged abuse. In particular, the narrow construction of the meeting competition defense would be more comprehensible if that defense were to be implemented in a framework that more accurately assesses the anticompetitive impact of the disputed conduct.

B. ANALYTICAL FRAMEWORK

The precautionary principle is embodied in the analytical framework adopted by DG COMP. The Discussion Paper is devoted primarily to abuse, but equal attention should be given to the discussion of the prior issues (market definition and dominance) as well as the scope allowed for justification of conduct that is prima facie abusive because the impact of Article 82 is a function of the mutually reinforcing effect that the treatment of all four elements has.

Consistent with the precautionary principle, the Discussion Paper appears to proceed on the basis that:

- **(a)** markets should be narrowly defined;
- **(b)** dominance is principally a function of a firm's share of a (narrowly-defined) market;
- **(c)** abuse is strongly dependent on the assumption that that dominance

23 It does not follow that, where the dominant position is stronger, the possibility of entry should be disregarded. On the contrary, while greater skepticism may be justified in that case, the Commission should always be careful to consider realistically the possibilities available to putative entrants.

entails a weakened state of competition that creates a need for active intervention; and

(d) justification must be stringently defined and strictly applied to ensure that the abuse so identified is not inadvertently permitted to continue.

1. Market Definition and Dominance

If an analytical framework of that kind is adopted, it is critical that the connected issues of market definition and dominance are analyzed in a way that provides a sound framework for the identification of abuse. While DG COMP correctly and helpfully recognizes some technical issues,[24] its opening premise points to what appears to be an excessively narrow approach when it focuses on *actual* competitors that provide an *immediate* competitive constraint to the putative dominant firm.[25] Taken literally, this approach ignores less immediate, but nonetheless real, competitive constraints on a firm such as those presented by supply-side substitutes. Its effect is compounded by an approach to dominance that likewise focuses on static considerations.

> IT IS CRITICAL THAT THE CONNECTED ISSUES OF MARKET DEFINITION AND DOMINANCE ARE ANALYZED IN A WAY THAT PROVIDES A SOUND FRAMEWORK FOR THE IDENTIFICATION OF ABUSE.

While dominance is expressed (non-controversially) as the possession of "substantial market power," that proposition comes to be defined by the proxies that are used to measure it. Although DG COMP acknowledges the limitations of market shares and the need to explore wider competitive conditions (notably, the significance of entry barriers), in practice the primacy of market shares is maintained:

> It is *very likely* that very high market shares, which have been held for some time, indicate a dominant position. This would be the case where an undertaking holds 50% or more of the market, provided that rivals hold a much smaller share of the market. In the case of lower market shares, dominance is more likely to be found in the market share range of 40% - 50% than below 40%, although also undertakings with market shares below 40% could be considered to be in a dominant position. However, undertakings with market shares of no more than 25% are not likely to enjoy a (single) dominant position on the market concerned. (footnotes omitted and emphasis added)[26]

24 The significance of the marginal customer is a case in point.

25 Discussion Paper, *supra* note 1, at para. 12.

26 *Id.* at para. 31.

The discussion of entry and expansion focuses on the extreme ends of the spectrum. Thus, such barriers are likely to be found if previous attempts to enter or expand have been unsuccessful. At the opposite end of the spectrum, where entry has been frequent and successful, such barriers are not likely to be found. DG COMP leaves open its approach to those cases that fall between the two extremes. The tone of the document, however, suggests that it will favor a finding of dominance absent proof of successful entry.

Cumulatively, this discussion means that Article 82 is applicable in a range of widely differing circumstances where, at the lower end of the range, the existence of substantial market power is highly contentious. That only emphasizes the need, at a minimum, to develop rules that recognize the differences between those cases and to avoid a precautionary principle that is insensitive to such differences.

2. Abuse

The general discussion of abuse is almost wholly concerned with the concept of foreclosure and does not extend to other questions, such as the relevance of intent or sacrifice.[27] The level of foreclosure that engages Article 82 is a function of two components: the level of competitive harm and the degree of probability that that harm will occur. The Discussion Paper articulates a standard based on the likelihood of foreclosure where foreclosure includes eliminating, constraining, and disciplining competitors. By contrast, *Microsoft* articulated a standard based on a risk that competition would be eliminated.[28] The higher level of probability in DG COMP's proposed standard is balanced by a lower level of competitive harm.[29]

---

27 DG COMP quotes the hallowed formula originally stated in *Hoffman-La Roche* which distinguishes abusive behavior from "normal competitive behaviour." Case 85/76, Hoffmann-La Roche v. Commission, 1979 E.C.R. 461 [hereinafter *Hoffman-La Roche*]. In practice, that concept has had little impact on the application of Article 82. Normally, it is trumped by the other elements of the *Hoffman-La Roche* rule which has come to embody the dominant firm's special responsibility not to weaken competition. In the Discussion Paper, the concept appears through DG COMP's recognition that broad adoption of certain types of behavior indicates their efficiency-enhancing potential but that has little, if any, impact on the way in which the foreclosure and efficiency tests are applied to dominant firms' employment of such behavior.

28 *See, e.g.,* Commission Decision, Case COMP/C-3/37.792, Microsoft (Mar. 24, 2004, not yet reported) [hereinafter *Microsoft*], at para. 992.

29 Given the Commission's concern to preserve its position in relation to past cases and (especially it may be imagined pending litigation), it is most unlikely that any substantive change is intended. In that context, note that one of the FAQs accompanying the Discussion Paper (accessible through the reference at note 1) stated that:

> There is nothing in the discussion paper that calls into question any of the Commission's past decisions. At the same time, the Commission must always work to improve its decisions and its policies. The review is about a better focus and a better argumentation in future cases. Furthermore, the fact that if the discussion paper leads to a more refined economic analysis, the Commission would in future argue a case in a different way than in the past, does not mean that the decision taken in a past case was wrong, only that the argumentation would today have been different.

The opinion of the Advocate General in *Virgin/British Airways (ECJ)* appears to articulate both approaches, stating that what is to be proved is the "likelihood" that the disputed conduct will "hinder" the maintenance or development of competition and deducing from that a requirement to prove that the rebates offered by BA were "capable" of making it "difficult or impossible" for its rivals to have access to the market and its business partners to choose between various sources of supply.[30]

The precise content of both elements of the standard is, however, important. First, if it is decided that a high degree of competitive harm (such as elimination of competition) is required to engage Article 82 in a particular case, then it is important to avoid assigning a low level of probability because that would effectively reduce the observed level of competitive harm at which intervention occurs. To treat a risk of competitive harm as sufficient to justify intervention sacrifices the gains to consumer welfare that unfettered competition would create if that risk did not materialize.

Second, the required degree of probability has a substantial bearing on the nature and quality of the evidence to be produced. If it is sufficient to show that there is a risk of elimination, then it is easy to slip into reliance on weak evidence and speculative analysis. Yet, that takes us into an area where, as the European Court of Justice (ECJ) has said, "the chains of cause and effect are dimly discernible, uncertain and difficult to establish."[31] It is the consequent risk of a false prognosis that led the Court to insist that the Commission provide convincing evidence of its theories of harm. What holds true for merger control holds no less true in the context of Article 82.[32]

The two components of the standard need to be determined independently. The required degree of harm should be determined by reference to the conduct and market impact that would substantially lessen competition in that context. The required degree of probability should be a likelihood. Anything less lacks evidential rigor and will dilute the standard of harm.

3. Objective Justification

The strength of the precautionary principle is also demonstrated by the narrow scope given to the concept of objective justification. The so-called "efficiency defense" states four cumulative elements, reflecting those laid down in Article

30 *Virgin/British Airways (ECJ)*, AG Opinion, *supra* note 5, at para. 71.

31 Case C-12/03P, Commission v. Tetra Laval (Feb. 15, 2005, not yet reported), at para. 44.

32 The distinction between the categories of case should not be exaggerated. While merger control cases are wholly forward-looking, Article 82 cases may contain substantial forward-looking elements: *Microsoft*'s focus on the risk that competition will be eliminated is a strong example. In any event, whatever differences there may be do not justify a difference in the legal standard. It may simply mean that the legal standard is easier to satisfy in one case than the other.

81(3), which cumulate the requirements of indispensability and non-elimination of competition. DG COMP acknowledges explicitly that "ultimately the protection of rivalry and the competitive process is given priority over possible pro-competitive efficiency gains."[33]

Without necessarily limiting the scope of this constraint, it is clear that it will have the most direct impact in respect of those firms that might be regarded as "super-dominant." Yet, it is precisely because of such firms' super-dominance that consumers are particularly dependent on them for efficiency gains. Assuming that the conduct that would otherwise be abusive truly is indispensable to realizing those efficiencies, DG COMP's position entails the long-term denial of those gains to consumers.

The Discussion Paper itself appears to recognize an exception to this principle when it says that it is lawful for a dominant firm to withhold supplies of an essential input, at least for a sufficient period to recover its investment, even though that results in an elimination of competition for a period.[34] DG COMP could, of course, reply that there is no conflict with its basic principle because the entitlement to withhold supplies is always time-limited so that there can be no permanent elimination of competition.[35] There is, however, no connection between that limit (which is defined by reference to the dominant firm's payback period) and a limit based on the elimination of competition. A stronger form of the argument would say that the time limit must be curtailed even further if that is necessary to avoid the elimination of competition (for example, because withholding the input even for the duration of the minimum payback period would be sufficient to choke off subsequent entry). To accept that argument entails acceptance of the proposition that it is preferable to forego the benefit of that development, however valuable to consumers it might be, than to risk the elimination of competition in the supply of goods that are less valuable to consumers. The legislator can, of course, decide that that is the right choice, but it is a mistaken choice to make for all cases and under all circumstances.[36]

33 Discussion Paper, *supra* note 1, at para. 91.

34 *Id.* at para. 235.

35 The concept of a universal time limit on the entitlement to withhold supplies is highly controversial and is discussed in Section VI.D of this paper.

36 *Microsoft* indicates the Commission's apparent willingness to grasp that nettle. The Commission canvassed, only to reject, the possibility that Microsoft's incorporation of Windows Media Player (WMP) into the Windows operating system could be justified on the basis that there was a cognizable benefit in the certainty of a single platform standard. It said that: "Under Community competition law an undistorted competition process constitutes a value in itself as it generates efficiencies and creates a climate conducive to innovation (innovation being, in markets such as the software market, a key competition parameter)" (*see Microsoft, supra* note 28, at para. 969). Unfortunately, that statement in that context is rhetorical rather than substantive. The Commission concluded that incorporation of WMP into Windows was not necessary to accomplish the benefits of integration, but then more crucially, allowed Microsoft to continue selling that product provided that it also made available a

It should also be said (if the contrary is argued) that neither consistency with Article 81(3) nor Article 82's own purpose compel that outcome. While the two provisions share the same direct objective (preservation of a system of undistorted competition), that objective is a means to maximizing consumer welfare and must be seen within the context of the Treaty's primary aim to secure economic development. Consumer welfare or economic development may prevail over the competition rules (unless the Treaty expressly dictates otherwise as it does in the case of Article 81(3)) where conduct that would otherwise infringe those rules is necessary to secure one of those goals. There is, moreover, a substantial difference between controlling an agreement between independent undertakings and imposing an affirmative duty to supply on a single undertaking. While freedom of contract is an important value that is recognized and protected by EC law, both EC law and the individual legal systems of the EC Member States recognize a broad competence to regulate that freedom where that is appropriate in the interests of public policy. By comparison, an undertaking's freedom to choose its business partners and to dispose of its property as it chooses are fundamental values that European legal systems are reluctant to limit except in the most extreme circumstances. To put it simply, the law is more willing to tell a citizen that he may not conclude a certain contract than it is to fetter his property rights. Consistently, competition law should be more willing to apply Article 81 to prohibit a particular agreement than it should be to apply Article 82 to force a firm to deal with its property against its will.

The meeting competition defense poses a similar dilemma in view of the two restrictions imposed by DG COMP. First, it requires that the dominant firm's response be proportionate to the competitive challenge. The proportionality condition is expressed in the narrowest terms possible, demanding proof that the conduct is the least restrictive option available to the dominant firm and that it is pursued for the absolute minimum period of time.[37] Second, it requires that the response does not significantly delay or hamper competitive entry.[38] It is axiomatic that successful price competition must delay or hamper entry. The only question, therefore, is whether that effect is significant. It has to be said that it is hard to envisage what sort of decision a dominant firm can make faced with that problem. To say that it may compete provided that, by and large, it fails is not compelling either to the firm or to public policy.

footnote 36 cont'd
version of Windows that did not incorporate WMP. The Commission, therefore, did not have to confront the hard choice which, in principle, this issue provokes.

37 Discussion Paper, *supra* note 1, at paras. 81-83.

38 *Id.* at para. 132.

III. Predation

DG COMP defines predation by reference to the conventional concepts of a short-term sacrifice that causes foreclosure and is recouped thereafter.[39] The Discussion Paper recognizes the need to distinguish between price reductions that form part of the normal competitive process and predation. The litmus tests that it proposes follow a cost-based model developed from the existing precedents. In summary:

(a) prices below average avoidable costs (AAC) infringe in the absence of credible explanation;[40]

(b) prices between AAC and average total costs (ATC) infringe if a predatory strategy can be established; and

(c) prices above ATC only infringe in extreme circumstances.

The only significant change is the replacement of the average variable costs (AVC) standard stated in *AKZO*[41] by an AAC standard.

A. THE CONCEPT OF PREDATION: SACRIFICE AND RECOUPMENT

While DG COMP's general definition of predation is founded on sacrifice and recoupment, two complementary elements of seemingly equal standing, its discussion of the evidence for predation treats recoupment as a possible, but non-essential, element.[42] DG COMP finally concludes that, in general, proof of dominance is sufficient to establish the likelihood of recoupment.[43] So, the notion of recoupment appears to progress from an element that it is important to prove independently to an element that is assumed to exist by virtue of the proof of dominance.

DG COMP bases its assumption on a finding of dominance without considering the quality of that dominance and, in doing so, does not respect its own direction to determine whether the disputed conduct is likely to have an exclusionary effect in the specific circumstances of the case.[44] In a case where dominance co-exists with active (albeit not fully effective) competition and the prey

39 *Id.* at para. 93.

40 DG COMP gives the conventional example of a new product launch.

41 Case C-62/86, AKZO v. Commission, 1991 E.C.R. I-3359 (ECJ).

42 Discussion Paper, *supra* note 1, at para. 115.

43 *Id.* at para. 122. That is, in outcome, consistent with the position taken by the ECJ that it is unnecessary to prove recoupment because EC law does not have to wait until the predatory strategy has succeeded. *See* Case T-83/91, Tetra Pak v. Commission, 1994 E.C.R. II-755, 1997 4 C.M.L.R. 726 (CFI), and on appeal, Case C-334/94P, 1996 E.C.R. I-5951, 1997 4 C.M.L.R. 662 (ECJ).

44 Discussion Paper, *supra* note 1, at para. 22.

remains in the market, it cannot simply be assumed that the dominant firm will be able to recover the full amount of any sacrifice that it makes.

As a general proposition, placing such weight on the sacrifice element imposes a burden that is too great for it to bear, largely because it is so hard to identify with precision whether a sacrifice has been incurred. Consider, for example, the case where a dominant firm responds selectively to an entrant's introductory price. If the dominant firm's discount is no greater than the absolute minimum that is required to win the contract and the contract makes a contribution to the dominant firm's sunk costs, it is not evident that there is any sacrifice at all: the dominant firm is better off having entered the contract than it would have been had it not entered the contract at all. The line between sacrificial and non-sacrificial behavior is hard to discern with precision and it is for that reason that evidence of recoupment is needed to strengthen the analysis.

This analysis points to the conclusion that it is a mistake to rely exclusively on any one element. A finding of predation necessitates an assessment that considers the evidence relating to sacrifice, exclusion, and recoupment to determine whether the conduct does indeed display the characteristics of a predatory strategy correctly identified by DG COMP in its introduction to this section.

B. PRICES BELOW AAC

DG COMP's treatment of pricing below AAC as presumptively predatory is unobjectionable in principle, but it does conceal a number of challenging evidential questions. First, what costs should be treated as avoidable? DG COMP contents itself with saying that, while in many cases AAC will equate to AVC, in some cases it will exceed AVC.[45]

Second, over what period of time is the possibility of avoidance to be considered? DG COMP proposes to take the period over which the alleged predation has occurred or (if it is still continuing) the period over which it is expected to occur.[46] In the latter case, is the Commission entitled to assume a period as long as it takes to secure the (assumed) foreclosure effect? Or should a finite period be selected? What basis is there, absent specific evidence, for preferring either choice?

The difficulties with a general use of the AAC standard (which could point equally to over- or under-assessment of avoidable costs) are such that, as framed in the Discussion Paper, it provides an inadequate basis on which to proceed. As a general principle, it is important that competitive harm be proved rather than presumed. That is especially important here because the location of the AAC

45 The only specific example that it gives of the latter condition is where the dominant firm invests in excess capacity to allow it to predate. *See id.* at paras. 108 and 109.

46 *Id.* at para. 105.

boundary could determine the outcome. Uncertainty as to its location can only chill price competition. For those reasons, it is preferable to proceed on the basis of a rule that AAC are taken to be equal to AVC unless there is clear evidence that some additional costs should be included such as the investment in excess capacity identified by DG COMP.

C. PRICING BETWEEN AAC AND ATC

DG COMP identifies three grounds, any one of which is sufficient to establish a predatory strategy. They are:

(1) direct evidence of intent;

(2) the absence of any reasonable commercial explanation for the pricing; and

(3) other sufficient indirect evidence.[47]

The most acute problems occur with the third category. By definition, there is no direct evidence of predatory intent and the strategy is capable of making commercial sense independently of its predatory effect. The question, therefore, is what factors are sufficient to justify rejection of the reasonable commercial explanation in favor of a finding that the dominant firm has a predatory strategy. DG COMP says, first, that a foreclosure effect must be shown and adds that it is usually necessary to investigate additional elements. Given that any successful price competition likely will satisfy the broad standard of foreclosure proposed by DG COMP, the additional elements become critical.

> THE CHALLENGES WITH DEALING WITH THESE CASES SUGGESTS THAT THE COMMISSION SHOULD BE WARY OF APPLYING ARTICLE 82 TO THESE PRICES IN THE ABSENCE OF CLEAR EVIDENCE OF PREDATORY INTENT OR LACK OF COMMERCIAL RATIONALE.

For that purpose, DG COMP identifies a list of factors—none of which is said to be necessary, but some unspecified combination of which is sufficient.[48] Of the factors identified, some are likely to be established by reason of the low pricing itself (e.g., an actual or likely exclusionary effect, the scale, duration and continuity of the low pricing, and an ability to recoup the short-term losses all fall into that category).[49] Others may be relevant, but will not be present in every case and need not be decisive differentiators (e.g., incurring specific costs in order to expand capacity that enables the dominate company to react to entry, concurrent application of other exclu-

47 *Id.* at para. 111 et seq.

48 *Id.* at para. 115.

49 It is striking that exclusion and recoupment are two essential facts to be proved in a predation case yet they are listed as optional components in the determination of whether a predatory strategy exists.

sionary practices,[50] reputational effects in other markets,[51] and the prey's particular dependence on external financing). While DG COMP includes the counter-strategies available to the prey in its list, it is not clear what significance those strategies have in the assessment process.

Absent those issues, the decisive factor is likely to be the characterization of a dominant firm's selective response to competitive entry. The narrow scope offered to the meeting competition defense, under which any successful response likely falls outside the scope of the defense, suggests an intention to prohibit any selective response that falls below ATC. An exception may be made where it can be shown that there is no sacrifice involved though, even then, the difficulties of relying exclusively on sacrifice to identify infringement discussed earlier in this paper should be recalled.

In any event, that outcome is a strong example of the dangers of the precautionary principle in a situation in which a firm, albeit dominant, confronts a number of rivals already established in the market. In such a case, price discrimination and pricing below ATC may be perfectly rational responses independently of any exclusionary effect. It is not obvious in those circumstances where the line between selective discounting that represents normal competitive behavior and selective discounting that represents predation should be drawn. While the change in behavior implicit in the concept of selective discounting may be relevant, the fact that behavior changes may simply reflect a competitive norm that could be lost through an overly extensive enforcement of predation rules.[52]

The challenges with dealing with these cases suggests that the Commission should be wary of applying Article 82 to these prices in the absence of clear evidence of predatory intent or lack of commercial rationale. When conduct has a plausible non-

50 Reliance on the cumulative effect of disparate pieces of conduct to establish that each is an abuse requires great caution to avoid the errors of the "monopoly broth" argument now generally repudiated under U.S. law.

51 The theory that a predatory reputation may have an exclusionary effect confronts the difficulty that the issue of predation typically arises following unsuccessful entry by a firm that was (presumably) not deterred by that reputation. In that case, as in most other cases of alleged predation, the entrant will already have incurred any sunk costs and it must be questionable, therefore, whether it is appropriate to set a rule that assumes that those costs will be recovered. It is also important to consider whether theories based on a reputation for exclusion are in fact well-founded. Paradoxically, the greater the number of instances where entry allegedly has been deterred, the more acute that requirement is-precisely because the more common the attempts at entry, the more questionable the deterrent effect of the reputation.

52 It may be relevant to distinguish between discounts offered to retain existing customers and discounts offered to attract new customers away from the entrant. In the former case, the dominant firm's incentive to meet the threat of competitive entry is obvious and, so long as it makes a positive contribution to fixed costs, seemingly reasonable. The latter case may be more questionable but, even here, it would be useful to compare the dominant firm's behavior on this occasion with its past conduct. If its present conduct is materially more aggressive than on past occasions, that may raise legitimate questions about why it is more concerned to win that business on this occasion. Even that, however, cannot be conclusive: firms' circumstances change and what may not have been a sensible discount in the past could have become one.

predatory explanation, that explanation should prevail unless the Commission can offer convincing evidence and argument to show that it is ill-founded. In such circumstances, it is questionable whether there is (or should be) any difference between these cases and those that fall within categories (1) and/or (2).

D. PRICING ABOVE ATC

DG COMP's identification of predation in this context is heavily influenced by its existing jurisprudence, most notably *Compagnie Maritime Belge*.[53] It does, however, take a broader position that predation may occur where the dominant firm enjoys non-replicable advantages or there are substantial economies of scale such that the dominant firm could price above its ATC and still exceed the entrant's ATC. While DG COMP seeks to limit the scope of this exception,[54] the concerns expressed in relation to a finding of predation in cases where prices fall between AAC and ATC apply with even greater force here. If it is ever acceptable to treat any price exceeding ATC as predatory, then it can only be treated as such in the most exceptional circumstances where there is incontrovertible evidence that it will lead to the creation or maintenance of absolute and persistent monopoly.

IV. Rebates[55]

A. DG COMP'S APPROACH

1. Overview

It was the case law on conditional rebates[56] that, above all, sparked the review of Article 82 and it is this area that shows the most innovative thinking by DG COMP. It no longer treats a conditional rebate as the functional equivalent of

53 Joined Cases T-24, 25, 26 and 28/93, Compagnie Maritime Belge v. Commission, 1996 E.C.R. II-1201, 1997 4 C.M.L.R. 273 (CFI), and on appeal, Joined Cases 395/96P et al., 2000 E.C.R. I-1365, 2000 4 C.M.L.R. 1076 (ECJ).

54 It states that there must be a clear strategy to exclude on the part of the dominant firm. The entrant must only be less efficient by reason of the non-replicable advantages or economies of scale and there must be specific price cuts that have the effect of deterring and preventing entry (Discussion Paper, *supra* note 1, at para. 129).

55 This topic has been the subject of extensive academic debate in recent years. *See* J. Kallaugher & B. Sher, *Rebates Revisited: Anti-competitive effects and exclusionary abuse under Article 82*, 25(5) Eur. Competition L. Rev. 263-285 (2004) and RBB Economics, Selective Price Cuts and Fidelity Rebates (U.K. Office of Fair Trading, Economic Discussion Paper, Jul. 2005. For an extended discussion, see the papers presented at European University Institute's Eighth Annual EU Competition Law and Policy Workshop, *available at* http://www.iue.it/RSCAS/Research/Competition/2003(papers).shtml and Symposium, *A Symposium on Loyalty Rebates*, 1(2) Competition Pol'y Int'l 89 (2005).

56 The Discussion Paper distinguishes between rebates that are conditional on the purchaser's buying behavior (such as purchasing a definable quantity of goods from the dominant firm) and those that are unconditional (such as those that are offered in respect of all the purchases made by selected customers). This paper is concerned solely with DG COMP's treatment of conditional rebates.

an exclusive dealing contract, recognizing that theories of predation have something relevant to say about the topic.

The essence of DG COMP's theory is that a rebated price is abusive (absent justification) unless either it covers the dominant firm's ATC or there is no evidence of possible foreclosure.[57] In applying that theory, DG COMP devises different price/cost models for retrospective rebates[58] and prospective rebates.[59] In both cases, narrowly circumscribed efficiency defenses are envisaged, but a meeting competition defense is ruled out.

It should be noted that the Advocate General in *Virgin/British Airways (ECJ)* takes no account of DG COMP's new thinking, saying that that case has to be decided under current legal standards.[60] In that context, the Advocate General says that, while the classes of exclusionary rebates are not closed, such an effect is to be expected in the normal course of events where targets are individually defined and retrospective rebates are employed.[61]

2. Retrospective Rebates: The Price/Cost Standard

DG COMP proposes a standard under which the rebate should fail if the share of the market at which the rebated price covers the dominant firm's ATC is greater than the share that an efficient entrant can reasonably be expected to capture.[62] That standard is explained by reference to a simple rebate system where there is a single threshold above which purchases qualify for a rebate on a

57 In the case of retrospective rebates, that occurs where (a) the rebate scheme does not affect a substantial part of market demand, (b) the threshold is set substantially below the level that customers would expect to buy from the dominant firm in any event, or (c) there are clear indications of a lack of foreclosure effect such as aggressive and significant entry and/or expansion by customers and/or switching of customers. In the case of prospective rebates, DG COMP only articulates the first of those possibilities.

58 Retrospective rebates are rebates that apply to the totality of a customer's purchases once a certain threshold has been passed.

59 Prospective rebates are rebates that only apply to the portion of the customer's purchases that exceed the threshold.

60 *See Virgin/British Airways (ECJ)*, AG Opinion, *supra* note 5, at para. 28.

61 *Id.* at paras. 47 et seq. The Advocate General also refers to the defendant's dominance, but that is necessarily present in any event (*id.* at para. 52). Although the Discussion Paper advances DG COMP's thinking in this area, it remains the case that under its new standards, retrospective rebates based on individual sales targets would only escape prohibition in exceptional cases. The Advocate General's opinion does not, therefore, create a material obstacle to the evolution of the law in the way contemplated by DG COMP.

62 Assuming a standard progressive rebate schedule (such that, as the volume of rebated sales increases, the marginal price declines), this test implies that the dominant firm's effective discounted price does not exceed its average total costs.

single scale.[63] The litmus test stated by the model is based on a comparison of two market shares:

(1) The required share, that is the share of the market at which the rebated price covers the dominant firm's ATC where:

- **a)** The rebated price is calculated on the basis that the entirety of the rebate is allocated to the sales that comprise the required share; and
- **b)** The ATC are calculated on the basis of a volume equal to the threshold specified in the rebate scheme.

(2) The commercially viable share, that is the share of the market that an efficient entrant can reasonably be expected to capture.

Having established those two shares, the test is disarmingly simple. If the commercially viable share exceeds the required share, the rebate scheme is non-exclusionary. Conversely, if the required share exceeds the commercially viable share, the rebate scheme is exclusionary.[64]

3. Prospective Rebates: The Price/Cost Standard

In the case of prospective rebates, DG COMP proposes a more straight-forward application of its predation standards under which a rebate should fail if the rebated price for purchases above the threshold does not cover the dominant firm's ATC.[65]

B. COMMENTARY

This commentary is organized in the following way. First, it sets out some reasons why DG COMP's approach to rebates, of both forms, is unduly restrictive. It follows with a consideration of two topics that present particular difficulty, namely the selection of a benchmark based on ATC and the treatment of retrospective rebates. Finally, it advances an alternative approach to that proposed in the Discussion Paper.

1. DG COMP's Overall Approach Is Unduly Restrictive

DG COMP acknowledges that rebates have an ambivalent effect, with the capability both to enhance efficiencies and to foreclose competitive entry and expan-

63 DG COMP does not elaborate on the application of its model to multi-tier rebate structures. As discussed in this paper, the differences between retrospective and prospective rebates once the initial threshold has been exceeded may not be large (*see* Section IV.B.3 of this paper).

64 Discussion Paper, *supra* note 1, at paras. 155 and 156.

65 *Id.* at paras. 166-169.

sion.[66] Consistent with its general approach, however, it applies its precautionary principle to the assessment of rebates. That approach is mistaken for two reasons: first, it takes an unduly narrow view of the pro-competitive function of rebates and, second, it takes an unduly optimistic view of the effect that regulation may have.

DG COMP's approach ignores two connected factors: first, conditional rebates may be an element in price competition and, second, rebates may achieve efficiencies that are no less real even though they do not approach the level of specificity or proof demanded by DG COMP. Those propositions are evidenced by the behavior of firms that commonly employ conditional rebates (of either form) even though they lack market power. In some cases, the starting point is a disagreement over price that is resolved by the use of a conditional rebate in what is a simple commercial deal trading volume for price. That rebate is quite likely to be a retrospective rebate, reflecting the fact that the deal is essentially one about the price for the totality of the supplies. From the supplier's perspective, the justifications for the rebate no doubt include the efficiencies that it derives from a strengthened expectation that the threshold volume of sales will be accomplished in the broad sense that likely capacity utilization over an extended period is increased.[67] There is no reason to believe that those factors are systematically inapplicable to dominant firms' employment of rebates. Similarly, in a distribution context, rebates provide a sales incentive that is useful to dominant and non-dominant firms alike.[68] It is, therefore, a mistake to proceed on the basis that a strict abuse standard has no adverse welfare consequences.

The strictness of that standard should be contrasted with DG COMP's faith in the price neutrality of regulatory intervention that is apparent in its statement: "The customer may not derive a direct benefit from the rebate system as the rebate may only bring the average price down to the level existing without the rebate system."[69] If that is intended to be a general proposition, then it is optimistic as an assessment of a dominant firm's likely behavior and mistaken as to the technical structure of prices under the two scenarios.

It assumes that the dominant firm would prefer to sacrifice the higher margin that it obtains on the assured sales for the prospect of a volume of contestable sales equal to that achieved with the rebate. That assumption fails, at least as a

66 *Id.* at para. 138.

67 A parochial example is provided by the experience of law firms where conditional price deals of the kind described in the text are not uncommon. Law firms' principal costs (staff, premises, and technology) are effectively fixed over a longer period than demand. There is, therefore, significant value in a pricing structure that gives greater assurance (though rarely certainty) of order flow.

68 For example, in *Virgin/British Airways (CFI)*, BA's competitors also offered incentive commissions of the kind employed by BA.

69 Discussion Paper, *supra* note 1, at para. 154.

general proposition, because the dominant firm may rationally prefer to retain the higher margin, especially where (absent the rebate) there must be a reduced expectation that it will achieve the same level of sales. DG COMP cannot consistently maintain that the rebate is objectionable because it induces increased sales and that removal of the rebate has no effect on such sales. In cases where that assumption does fail, there must be a probability that the price to the assured base will be held constant or, at least, not reduced to the average level produced by the rebate and that the effective price paid in respect of the contestable sales will rise.

More technically, the structure of the rebate systems that DG COMP opposes is such that their prohibition must reduce the intensity of price competition, at least in the short run. As DG COMP notes, the marginal price charged by the dominant firm is substantially below its average price.[70] It follows that the marginal price for the next unit that is not sold by the dominant firm will be even lower. As it is that price that a competitor must beat in order to secure a sale, even if the average price remains constant, raising the marginal price to equal the average price must raise the price that the competitor has to beat.

This is not to argue that rebates cannot have a foreclosure effect or that the Commission should not apply Article 82 to such cases. However, the balance of benefit and harm posed by rebates is far more complex than DG COMP allows and the precautionary principle that it has applied risks a loss of consumer welfare that may not necessarily be compensated by the long-term structural changes that it seeks to promote.

2. Universal Application of an ATC-based Standard Is Inappropriate

While DG COMP correctly stresses the importance of using predation theory to assess rebates, its implementation of that theory in this context requires that the dominant firm's effective price for sales that are, or should be, contestable must exceed the firm's ATC. By comparison, other commentators have proposed a test based on the dominant firm's AAC.[71] DG COMP justifies its departure from normal predation theory on the basis that, because the rebate structure is self-sustaining over the long term, it does not involve any sacrifice on the part of the dominant firm.[72]

70 *Id.* at para. 153.

71 *See* R. O´Donoghue, *Over-Regulating Lower Prices: Time for a Rethink on Pricing Abuses under Article 82 EC*, and D. Ridyard, *Article 82 Price Abuses - Towards a More Economic Approach*, presented at European University Institute's Eighth Annual EU Competition Law and Policy Workshop, *available at* http://www.iue.it/RSCAS/Research/Competition/2003(papers).shtml. *See also* H. Hovenkamp, *Discounts and Exclusions*, University of Iowa Legal Studies Research Paper, Number 05-18 (August, 2005).

72 Discussion Paper, *supra* note 1, at para. 154.

In taking that approach, however, DG COMP loses sight of two things. First, its analysis of predation correctly recognizes that it is necessary to fashion a rule that avoids undue constraints on price competition by dominant firms. DG COMP's approach to this issue neglects the fact that rebates are commonly a form of price competition and in effect treats the attachment of purchase conditions to a rebate program as conclusive proof of a strategy to predate. Second, the sacrifice analysis is a part of an analytical framework that is useful to differentiate 'good' from 'bad' price competition in certain situations. It does not follow from its inapplicability in other situations that a full cost recovery benchmark is dictated in those cases. The challenge remains to consider whether, as a matter of general principle and in the specific circumstances at hand, a rebate structure that produces an effective price below ATC should be held to have an anticompetitive effect.

So far as general principle is concerned, DG COMP's proposed test would capture discounts that are perfectly rational for any firm (dominant or not) to employ, independently of any exclusionary effect, thus weakening the intensity of competition already existing within the market. Given that state of affairs, Article 82 should not be applied in a way that is predisposed towards the prohibition of all conditional rebates that yield an effective incremental price below ATC. Furthermore, consistent with the approach to predation that should be adopted, it should be incumbent on the competition authority or the plaintiff to produce additional evidence and analysis that is sufficiently convincing to show that a substantial lessening of competition is likely to occur where the effective price exceeds AAC.

It should be incumbent on the competition authority or the plaintiff to produce additional evidence and analysis that is sufficiently convincing to show that a substantial lessening of competition is likely to occur where the effective price exceeds AAC.

3. The Treatment of Retrospective Rebates Is Unduly Hostile

The intensity of the attention that the Discussion Paper devotes to retrospective rebates suggests that if there is one way in which DG COMP wishes to influence dominant firms' behavior, it is to abandon such rebates. If there is any justification for a conditional rebate, it appears to say, it will be satisfied by a prospective rebate. Such a position, if it were to be intended, would ignore the pro-competitive value that such rebates can have. Even if the Discussion Paper does not intend to go that far, it overstates the differences between the two types of rebate and, as a result, proposes a regime for retrospective rebates that is unduly rigorous.

At a conceptual level, the difference between the effects that retrospective and prospective rebates are capable of having is insignificant. A simple example, set out in the following table, illustrates the point.

Figure 1

Level	Sales (£m)	Rebate (£K)	Prospective rebate %	Prospective rebate Per unit (£)	Retrospective rebate %	Retrospective rebate Per unit (£)
0	< 1.0	0	–	–	–	–
1	1.0 – 1.099	5	5	4	.45	4.5K/0.36
2	1.1 – 1.199	15	10	8	1.25	13.75K/1.00
3	1.2 – 1.299	30	15	12	2.3	27.6K/1.84

Note: Calculations assume an unrebated unit price of £80. The rebate in column 3 represents the total rebate payable for sales within the corresponding range stated in column 2. The unit values for retrospective rebate state, first, the value payable for the sale at the change of level and, second, the value payable for other sales within the level.

The example is, of course, oversimplified, but it does illustrate two obvious yet important points. First, the difference between the two types of rebate is a function of system design not inherent concept. Therefore, care should be taken to avoid rules that are based on the latter without allowing for consideration of the former. Second, the incentive created by a retrospective rebate is uneven while the incentive created by a prospective rebate is more consistent. Although a retrospective rebate clearly creates stronger incentives at the thresholds, a prospective rebate creates stronger incentives for sales between the thresholds. Which of those is the more important is once again a matter of design in the case at hand, not concept, especially when dealing with second and subsequent levels in multi-level schemes.

That said, the potential to design schemes where that impact is substantial means that a model that seeks to determine the effect of the scheme on sales below the initial threshold may be useful. In broad concept, the model proposed by DG COMP is not unreasonable, but there are three substantial areas of difficulty:

(1) It adopts a price/cost benchmark based on ATC that, for the reasons that already have been discussed in this paper, are fundamentally mistaken.

(2) The test is dependent on identifying the commercially viable share, that is the share of the market that an efficient entrant can reasonably be expected to capture. DG COMP states that it initially will base its assessment on the position of a competitor who wants to enter at minimum efficient scale.[73] However, that operates only as a measure of efficiency, not entry. We are left with a completely open question as to the level of entry that such a competitor should be assumed to achieve. That problem is made more acute by the absence of any guid-

73 *Id.* at para. 157.

ance as to the period over which a competitor's entry and expansion should be assessed. The focus is, moreover, on structural change rather than the rivalry that may occur within a given structure.

In the aggregate, the document points to a commercially viable share that is to the lower end of the plausible range. The significance of that lies in the fact that the lower the commercially viable share, the more likely it is (all else equal) that the rebate will fail the test. That, coupled with the fact that the test is based on a decision maker's assessment of likely performance rather than a cost-based assessment of comparable efficiency, means that the test provides dominant firms with a limited and unreliable basis on which to determine their behavior.

(3) DG COMP adds the rider that a retrospective rebate will be presumed to deter switching and so enhance loyalty where the rebate is either based on a percentage of customers' total requirements, is an individualized volume target, or is a standardized volume target where the thresholds are well targeted to customers' purchasing requirements.[74] It is unclear whether these features (which of course correspond to those found in the precedents that DG COMP is anxious to preserve) supplement or replace the price/cost model. Probably, they will be taken into account as a ground for (further) reducing the commercially viable share. No indication is given, however, as to the extent to which that will happen. Of course, it may be said that that is impossible outside the specific facts of a particular case but, conversely, that makes it impossible for a dominant firm to predict the likely application of the test and further diminishes its utility as a source of guidance for courts, agencies, or firms.

THE ASSESSMENT OF REBATES OF ANY FORM SHOULD BE BASED ON AN OVERALL CONSIDERATION OF WHETHER THEY SIGNIFICANTLY LESSEN COMPETITION BY COMPARISON WITH THE POSITION THAT WOULD BE CREATED FOLLOWING REGULATORY INTERVENTION.

4. An Alternative Approach[75]

Consistent with the approach advocated above, the assessment of rebates of any form should be based on an overall consideration of whether they significantly lessen competition by comparison with the position that would be created following regulatory intervention. While that involves an assessment of substantially the same issues discussed by DG COMP, it requires a fundamentally different approach to their content and application.

74 *Id.* at paras. 158 and 159.

75 This paper does not claim striking (or indeed any) originality for any sensible features of this proposal. There is, for example, a parallel discussion in RBB Economics' study (*see supra* note 55) which overlaps significantly, though not wholly, with this discussion.

The starting point is identification of an appropriate price/cost benchmark. For the reasons already stated in this paper, DG COMP's use of a benchmark based on ATC is mistaken. Use of that benchmark is not appropriate unless it is clear that the dominant firm's ability to charge an effective rebated price that is below ATC does indeed have an exclusionary effect and that requiring the dominant firm to raise that price to cover ATC would not lessen the intensity of competition currently prevailing in the market. That would not be expected to occur unless the dominant firm enjoys a virtual monopoly and it is evident that the rebate structure does have a material deterrent effect on entry.

When dealing with facts where there are established competitors in the market, the benchmark should not be lower than the dominant firm's AAC over the range of sales affected by the rebate scheme and should not exceed the AAC over the same range of a competitor that has the scale of the next largest firm and that is otherwise as efficient as the dominant firm.[76] The reason for selecting the upper bound is that rational competitors would disregard the sunk costs that fall between AAC and ATC in their pricing decisions and, therefore, would be willing to lower prices to that level. To set a threshold at a level that is higher than that would result in an immediate lessening of competition that would only be justified if there were strong grounds for believing:

(a) that encouraging further entry or expansion would eventually produce a market structure in which competition is more intensive than it is otherwise likely to be and

(b) that longer term gain outweighs the short-term loss of competition.

It seems that those requirements would rarely, if ever, be satisfied where the present facts are consistent with single firm dominance. They might be satisfied where the present facts reflect collective dominance, but that requires a level of analysis that the Discussion Paper does not attempt.

The merit of the lower bound is that it is, self-evidently, consistent with the as efficient competitor standard. It would capture any incremental fixed costs to be incurred by the dominant firm while, conversely, making proper allowance for any economies of scale or scope. The fact that it would disregard the greater investments potentially required of a smaller competitor in order to achieve a similar expansion is implicit in the notion of the as efficient competitor standard.[77] Many commentators would likely say that that should be dispositive of the issue. For them, there are no circumstances in which a rebate scheme that

76 In this context, the selection of AAC has greater merit than in the case of conventional predation because the typical rebate scheme is likely to be a much longer-run strategy than the paradigm cases of predation. That does not mean that the use of an AAC benchmark in this context is wholly free from the difficulties discussed in the context of predation, but they are not so substantial as to compel selection of an alternative benchmark.

77 See the discussion of the nascent "as efficient" competitor in Section II.A.2 of this paper.

satisfies that standard should be challenged and, even if there are risks that competitive entry or expansion would be constrained, they are outweighed by maximizing the retained level of price competition and by minimizing the incidence of false positives that would otherwise occur.

Even if that goes too far, at a minimum, those factors mean that strong grounds should be required to intervene where the rebated price does exceed the lower bound provided by the dominant firm's AAC. It would be necessary to identify evidence that demonstrates that an effective price between the upper and lower bounds will constrain the growth of effective competition by established or new competitors and that the design of the scheme indicates the pursuit of a strategy to exclude rather than the pursuit of any legitimate commercial objective. Among the factors that should be considered in that respect are:

(1) *The scale of the divergence between the upper and lower bounds*. Further inquiry should only be undertaken when the divergence is substantial because it is only in those circumstances that a rebate scheme falling into that range can have the significantly constraining effect on competition that is necessary to justify intervention.[78]

(2) *Actual evidence of the scheme's impact on entry and expansion by rivals*. DG COMP acknowledges that that is a relevant factor but only, it seems, if it can be regarded as aggressive and significant. Certainly, it has to be assumed, both from the language of the Discussion Paper and from the express endorsement of both *Michelin II* and *Virgin/British Airways (CFI)* (where evidence of market share loss was advanced without success), that it will not prevail unless there is evidence of a substantially greater loss of position than either of those companies suffered. Indeed, it might be said that the loss envisaged is so substantial that it would require a level and intensity of competition sufficient to cast doubt on a finding of dominance.[79]

78 Parenthetically, it may be said that the existence of a substantial gap between the two AAC values indicates that there are substantial efficiencies associated with the dominant firm's sales such that intervention creates a significant risk of promoting inefficient entry. That could lead to the proposition that either the efficiency gap is slight, such that the risk of successful foreclosure is too small to justify intervention, or that the efficiency gap is large, such that the risk of inefficient entry is too large to justify intervention. While that may be so, it is not sufficiently persuasive to conclude the inquiry without further investigation of the issues of competitive effect. It does underscore, nonetheless, the need for that inquiry to be rigorous.

79 Any argument that evidence of actual entry or expansion contradicts foreclosure is always vulnerable to the riposte that, but for the conduct in dispute, that entry or expansion would have been even greater. So stated, that is an unanswerable proposition because there is no evidence that can be advanced to contradict it. Indeed, as soon as it is accepted that the rebate program secures the dominant firm an extra sale, the proposition is substantiated. At the same time, that renders the debate meaningless. The step forward that the Discussion Paper does take is to restore meaning to that debate. The question is now no longer an unanswerable counterfactual but a question of the level of entry that contradicts foreclosure.

That approach is overly demanding. Evidence of continuous market share loss or fluctuating market shares suffered by the dominant should be sufficient to show that a rebate scheme does not have an exclusionary effect.

(3) *Counter-strategies available to competitors*. This concept should be construed broadly to include the availability of alternative distribution channels as well as other commercial responses. DG COMP itself says that, in assessing likely foreclosure effects, "[the Commission] will also consider the possibilities of the existing and possible future competitors to curb and counter the fidelity enhancing potential of the dominant company's conduct."[80] When read with the remainder of the Discussion Paper, however, that consideration appears to be too limited and undemanding. Where the answer lies in any particular case requires an assessment of its specific facts in which actual market evidence must be more compelling than prediction. In assessing that evidence, close attention should be paid to the fact that the dominant firm's rivals have managed to establish themselves in the market. If that has occurred despite the existence of the rebate scheme, then the Commission should approach claims that further expansion is constrained with a healthy measure of skepticism.

If, and only if,

(a) the scale of the divergence between the upper and lower bounds is substantial,

(b) the dominant firm's market share is stable or increasing, and

(c) there is insufficient evidence of available circumvention strategies

is it then necessary to consider whether the rebate scheme is predatory. Each of the specific factors enumerated below could be perceived as evidence of such a strategy.

(4) *The use of a retrospective rebate*. Although DG COMP's treatment of retrospective rebates is unduly harsh and draws an excessively sharp distinction between them and prospective rebates, it is equally not possible to exclude the fact that a retrospective rebate can have a particularly strong effect at the initial threshold.[81] The larger the value of the initial rebate relative to the customer's total purchases, the more

80 Discussion Paper, *supra* note 1, at para. 144.

81 As noted above, one cannot generalize about the relative effects of retrospective and prospective rebates at subsequent levels in a multi-tier scheme. The retrospective rebate's stronger effect at the thresholds is counterbalanced by the prospective rebate's stronger effect between the thresholds. Their relative potency in a specific case depends on factors such as the relationship between average order size and the scale of the levels. The closer together the two are, the more plausible it is to say that a retrospective rebate continues to exercise a suction effect throughout the range of the rebate scale. Conversely, the further apart they are, the less plausible such a claim is.

plausible it is to think that the rebate structure might be acting as a constraint. In those circumstances, it would be necessary to test the initial rebate's price/cost relationship. For that purpose, a structure similar to that devised by DG COMP would probably be necessary subject to modifications to:

(a) substitute AAC for ATC[82] and

(b) clarify the assessment of a commercially viable share.[83]

(5) *The degree of correlation between the initial threshold and increments in efficiency*. It would be a contrary indicator if the thresholds (especially the initial threshold in a retrospective rebate scheme) were set at a level that manifestly bears no relationship with the increments in efficiency (such as economies of scale and scope that partially explain the employment of rebates), especially if those increments occur at sales levels significantly lower than the thresholds. That factor would tend to confirm that the rebate was designed to exploit the higher assured base of sales that the dominant firm enjoys.

When considering this factor, it should be recalled that we only reach this point in the analysis where the divergence between the upper and lower bounds is substantial, implying that there are substantial economies to be achieved by growth to the scale of the dominant firm. It is perfectly possible, therefore, that an initial threshold that is set significantly above the level of sales achieved by rivals would be reasonable.

(6) *The format of the rebate*. At least in the context of retrospective rebates, DG COMP treats rebate structures based on percentage requirements, growth in purchases, individualized targets, and well-targeted standardized targets as presumptively loyalty-enhancing. That observation may be reasonable in a limited number of cases, where the targets directly or indirectly account for a high proportion of customers' requirements, but, otherwise, they appear insufficient to strengthen the case that the rebates have an exclusionary effect.

(7) *Other evidence of exclusionary intent*. Is there other credible evidence available to suggest that the rebate scheme has been designed to exclude or limit the growth of competitors rather than meet legitimate business objectives? That said, such evidence should be treated with skepticism if there is limited evidence of exclusionary intent that can be derived from points (4), (5), or (6) given that, if such intent exists, one would expect it to manifest itself in at least one of those ways.

---

82 Admittedly, application of an AAC-based model would not be easy and further work would be required to establish a robust approach.

83 So far as possible, the subjective elements in the structure proposed by DG COMP should be excluded.

> A conventional predation analysis should conclude by asking whether recoupment is plausible, normally through raised prices following foreclosure brought about by price reduction. In this context, where that initial sacrifice is not an essential feature of the strategy, the question manifests itself rather in the form of whether it enables the dominant firm concurrently to maintain prices at a level higher than they would be otherwise. Essentially, that becomes subsumed in the general question of whether the rebate substantially lessens competition. While the specific factors suggested above should inform that analysis, they cannot exclude the need for a full assessment of the effects of the conduct that is under scrutiny.

Finally, it is important to recall that this is only the first stage in the process. Even if the rebate scheme fails these tests, it is still open to the dominant firm to advance a legitimate business reason to justify the scheme. The Discussion Paper places severe constraints on the justifications that may be advanced. Not only does it do inadequate justice to the efficiency arguments that may be advanced, it wholly excludes the possibility of relying on a meeting competition defense. There is insufficient justification for such an approach. If competitors choose to use a rebate scheme as a competitive tool, it must be permissible (and desirable for consumer welfare) for a dominant firm to respond proportionately in the same way.[84]

V. Tying[85]

A. DG COMP'S APPROACH

DG COMP adopts a conventional approach to the definition of tying,[86] identifying the four usual elements, namely that:

(1) there are distinct tying and tied products;

(2) the firm concerned is dominant in the market for the tying product;

---

84 It is not a sufficient response to that observation to say that competitors' use of a rebate scheme may be an effort on their part to mitigate the dominant firm's scheme. The fact that they are able to use a scheme that is structurally comparable with that of the dominant firm tells us something important about the conditions of competition on that market.

85 For a recent discussion of this area, *see* Colloquium, *A Colloquy on Tying*, 1(1) Competition Pol'y Int'l 1 (2005). *See also* B. Nalebuff, *Bundling, Tying and Portfolio Effects*, DTI Economics Paper 1, Feb. 2003, *available at* http://www.dti.gov.uk/ccp/topics2/pdf2/bundle1.pdf and C. Ahlborn, D. Evans, & J. Padilla, *The Antitrust Economics of Tying: A Farewell to Per Se Illegality*, Antitrust Bull. (2004).

86 Tying is defined broadly to cover both contractual and technological tying as well as pure price bundling (where the products are only available at the bundled price) and mixed price bundling (where the bundled price offers a discount to the sum of the stand-alone prices at which the bundle components are offered).

(3) the tying practice is likely to have a market-distorting foreclosure effect; and

(4) there are no efficiency or other justifications for the tying practice.

The discussion of the (critical) third test concludes that:

> "Where the Commission . . . finds that the dominant company ties a sufficient part of the market, the Commission is likely to reach the rebuttable conclusion that the tying practice has a market distorting foreclosure effect and thus constitutes an abuse of dominant position."[87]

DG COMP envisages that a tie may be justified where it produces cost savings or an improved product provided that it satisfies the four general pre-conditions.

B. COMMENTARY

The Discussion Paper says little about this topic that is new, a fact that is unsurprising given that *Microsoft* is pending before the CFI. Therefore, it will not satisfy those critics who have said that EC law's predisposition to prohibit tying mischaracterizes a commercial practice that has a ubiquity (among dominant and non-dominant firms alike) that argues strongly that it is generally efficiency-enhancing and, therefore, merits a predisposition to permit.[88]

The Discussion Paper acknowledges at the outset that tying and bundling are common practices that often have no anticompetitive consequences for dominant or non-dominant firms alike.[89] The Discussion Paper correctly qualifies that general assessment by saying that tying can, in certain circumstances, lead to anticompetitive consequences, of which foreclosure is the only issue that is considered in this paper.[90] It then moves from that premise to articulate the general rule (quoted above) that, provided it has sufficient market coverage, a tying policy will give rise to a "rebuttable" presumption of abuse.

87 Discussion Paper, *supra* note 1, at para. 188.

88 For a detailed discussion of these issues, see the articles cited in *supra* note 85.

89 Discussion Paper, *supra* note 1, at para. 178.

90 *Id.* at para. 179. Price discrimination and price elevation are exploitative abuses and so fall outside the scope of this paper. They will be considered in the second phase of DG COMP's study.

Given DG COMP's own starting point, that is inappropriate. First, as a matter of general principle, the legal burden rests on the Commission to establish the essential elements of the abuse. In this context, DG COMP's Discussion Paper suggests that foreclosure arises in three specific situations.[91] If indeed they do define the circumstances in which tying may give rise to market foreclosure, then proof that one or more of those circumstances exists forms an essential part of establishing the abuse that must be discharged by the Commission.

Second, the adverse effects specified by DG COMP appear to be dependent on a more severe degree of foreclosure than is consistent with its general description of foreclosure, that is any limitation on a rival's entry, expansion, or competitive independence.

Third, the finding of foreclosure presupposes the existence of a tie. DG COMP's discussion of that topic requires development in the contexts of technological tying and mixed bundling. It presupposes that the incorporation of one product into another necessarily creates a tie between those products without addressing the question whether that has the quality of coercion that is the hallmark of a hard tying practice.[92] As for mixed bundling, it stipulates a standard for the pricing of the tied component of the bundle based on the long run incremental costs associated with that product.[93] That approach is equivalent to the use of an ATC-based standard for rebates and is open to objections comparable with those discussed in that context.

Fourth, the generally benign assessment of tying sits uneasily with the very limited circumstances under which DG COMP contemplates that a tie may be justified. That problem, of course, is particularly acute if the law states an overly broad prohibition (although the correct solution is to restate the scope of the prohibition). Independently of that issue, a defense based on necessity must discount valuable efficiency gains. This is particularly relevant to the case of product integration.

More specifically, the Discussion Paper's approach to efficiencies ignores metering as a non-exclusionary commercial explanation for tying. Admittedly, that may raise issues of exploitation (both as to price discrimination and price elevation) which will require consideration in that part of DG COMP's analysis.

91 Namely, where (a) foreclosure in relation to the tied product is caused by a reduction in the number of potential customers available to firms that only supply the tied product, leading to their marginalization, exit, or restricted entry, or (b) foreclosure in relation to the tying product is caused by (i) making it impossible to supply the tying product without also supplying the tied product or (ii) causing exit in respect of a tied product that could eventually compete with the tying product.

92 This issue is central in *Microsoft* where the company argues that consumers' ability to download alternative media players at no cost means that the integration of Windows Media player into the Windows operating system lacks any coercive element.

93 Discussion Paper, *supra* note 1, at para. 190.

However, it would be wrong to preempt that debate by fashioning a rule that treats tying in such circumstances as an unjustifiable instance of foreclosure.

DG COMP's treatment of the transmission of consumer benefits is obscure. It says that tying should be considered abusive when a retailer is able to obtain, on a regular basis, supplies of equivalent products on the same or better terms than those offered by the supplier that applies the tying practice, as evidently the pass on is not realized. In the first instance, it is necessary to consider how, in those circumstances, the tie creates a market foreclosure that requires justification. If the relevant products (it is not clear whether tying and tied products or tied products alone are intended) are available on the same or better terms, one would expect customers to buy them. Second, that caveat reduces consumer benefit to a matter of price or other contractual terms. There may well be other non-contractual benefits that customers derive from the tie (such as the convenience of the combination for which they may be willing to pay more) that are cognizable under this standard. Indeed, the fact that customers are willing to pay more for a tied product when the untied products are freely available may simply confirm the value of the tie to customers, not its harmful effects.

The fact that customers are willing to pay more for a tied product when the untied products are freely available may simply confirm the value of the tie to customers, not its harmful effects.

VI. Refusal to Supply[94]

A. DG COMP'S APPROACH

DG COMP's discussion of refusal to supply broadly reflects the existing jurisprudence of the EC Courts and the Commission's own decisional practice. It differentiates three principal categories. In all three cases, it is necessary to show that the refusal is likely to have a negative effect on competition and the refusal may be excused by an objective justification. The three cases are differentiated in that:

(1) where the refusal concerns the termination of supplies to an existing customer, no additional element needs to be established;

(2) where the refusal concerns the refusal to start supplying an input, it is also necessary to establish that that input is indispensable to a firm's ability to carry on normal economic activity in a downstream market; and

94 For a recent discussion of these issues, *see* Symposium, *Aspen Skiing 20 Years Later,* 73 Antitrust L.J. 59 (2005). *See also* D. Geradin, *Limiting the scope of Article 82 EC: What can the EU learn from the U.S. Supreme Court's judgment in Trinko in the wake of Microsoft, IMS and Deutsche Telekom?*, 41 Common Market L. Rev. 1481-1518 (2004) and C. Ahlborn, D. Evans, & J. Padilla, *The Logic & Limits of the Exceptional Circumstances Test in Magill and IMS Health*, 28 Fordham Int'l L.J. 1109-1156 (2005).

(3) where the refusal concerns the refusal to license an intellectual property right, it is also necessary to establish that:

(a) the IPR is indispensable (as described in the previous case); and

(b) the would-be licensee intends to produce new goods or services for which there is a potential consumer demand and that are not simply duplicative of those supplied by the dominant firm already.

At the end, DG COMP tacks on two brief and somewhat opaque paragraphs to deal with the refusal to supply interoperability information that clearly have been included with *Microsoft* in mind. They state that, while a dominant firm generally is not obliged to secure interoperability between one market and another, it may be an abuse to use the refusal of interoperability information to leverage market power from the dominated market into another market and, in those circumstances, lower intervention thresholds (for example, with regard to the protection of trade secrets) may be justified.[95]

B. THE CONCEPT OF A MARKET IN THE CONTEXT OF REFUSAL TO SUPPLY CASES

Refusal to supply doctrines are commonly described, as they are by the Discussion Paper, in terms of upstream and downstream markets. The way in which the concept of a market is defined for this purpose determines the circumstances in which a refusal to supply can be challenged. In straightforward cases, such as those involving a termination of existing supplies, the fact of the existing commercial relationship answers the question of whether there are two separate markets.

Cases where the complaint is that such relationships do not exist when they should raise potentially more complex questions about the delineation of markets. DG COMP restates the answer provided by the ECJ in *IMS Health* to the effect that, for this purpose, "it is sufficient that a potential market or even hypothetical market can be identified."[96] While the Discussion Paper simply echoes existing EC law, it is important to appreciate the implications of that ruling for cases where new supplies are demanded. Although the Discussion Paper continues to use the terminology of vertical market relationships, it is clear that Article 82 can be applied to require the supply of an input by one competitor to another horizontal competitor where the requirements of competitive impact, indispensability, and novel product are met.

C. ALL CASES: A NEGATIVE EFFECT ON COMPETITION

DG COMP states that satisfaction of this condition depends on the state of the pre-existing competition in the downstream market. Two specific cases are dis-

95 Discussion Paper, *supra* note 1, at paras. 241 and 242.

96 *Id.* at para. 227. *See IMS Health*, *supra* note 17, at para. 44.

cussed. In the first case, where the dominant firm is not present and there are several competitors in the downstream market, a negative effect is said to be unlikely unless the refusal to supply is likely to lead to collusion in that market. The second case concerns the situation where the dominant firm is present in the downstream market and there are few competitors present in that market. In that case, DG COMP draws a distinction between the termination of existing supplies and the refusal to commence supplies.[97] Where supplies are terminated, a negative effect on competition will normally be presumed. Where supplies are not commenced, DG COMP simply says that a negative effect is more likely than in the first case (where the dominant firm is not present and there are several competitors in the downstream market).

In no case is it necessary that competition should be completely eliminated in the downstream market. Beyond that, DG COMP does not specify in its assessment of the different classes of refusal what level of competitive impact is necessary to engage Article 82. In its earlier description of refusal to supply, it describes the exclusionary effect as exit, marginalization, or non-entry of the competitor to the downstream market and goes on to say that "[f]or a refusal to supply to be abusive, it must . . . have a likely anticompetitive effect on the market which is detrimental to consumer welfare."[98]

Although this test is common to all three forms of refusal to supply, it is likely that not only its application but also its content will depend on whether the refusal concerns a termination of existing supplies or a refusal to commence supplies. In the latter case, as the content of the test cannot be separated from the additional indispensability and new product tests, the discussion will be continued in the subsequent sections that address those tests.

In the case of termination, there is no additional requirement to be fulfilled. Nonetheless, DG COMP's assessment of foreclosure in such cases is unsatisfactory in a number of respects. First, it says that foreclosure will be presumed where supplies to one of the dominant firm's few competitors on that market are terminated. That statement conflicts with its own direction to consider the specific market impact of the disputed conduct.[99] Second, in any event, it is silent as to the number of competitors that constitutes "a few" for this purpose and the criteria by which that should be determined. Third, it does not address at all the case where several competitors remain in the downstream market.

As a practical matter, it is impossible to dissociate the assessment of the termination's competitive impact from the reasons for its occurrence. Where the termi-

97 Throughout this and the following sections, a refusal to commence supplies includes a refusal to grant an IP license.

98 Discussion Paper, *supra* note 1, at para. 210.

99 *Id.* at para. 22.

nation constitutes a disciplinary measure against a customer for pursuing, or threatening to pursue, a commercial policy that intensifies competition to the dominant firm, the requisite competitive effect likely will be established. That is understandable provided that a sufficient level of market impact is shown.[100] Conversely, where the termination arises from routine and non-contentious commercial factors (such as non-payment of bills), an adverse conclusion is unlikely.

The most challenging cases occur where the dominant firm changes its approach to the downstream market by, for example, vertically integrating or reorganizing its distribution system. The immediate response, that there must be a negative impact on competition, is too vague to advance the debate much further. In such cases, there should be clear additional evidence that the termination of supplies to that customer will distort competition by creating or reinforcing market dominance on either the upstream or downstream market. Market dominance in this context may be either single firm dominance or collective dominance. In the latter case, it would be necessary to show that the conditions for collective dominance are satisfied and that (by reason, for example, of the terminated customer's business model) termination of supplies to that customer creates or reinforces those conditions. Where several competitors remain in the downstream market, it is unlikely that a finding of dominance (in either form) would be justified.

D. REFUSAL TO COMMENCE SUPPLY OR LICENSE IPRS: INDISPENSABILITY

Indispensability is described in the following terms: "A facility is an indispensable input only when duplication of the existing facility is impossible or extremely difficult, or because a second facility is not economically viable in the sense that it would not generate enough revenue to cover its costs."[101] Although DG COMP bases that description on *Bronner*,[102] it is not absolutely clear that the concept of economic viability that it states is wholly consistent with *Bronner*. *Bronner* treats lack of economic viability as an instance of the underlying requirement of impossibility or extreme difficulty rather than an alternative explanation of indispensability (which, purely as a matter of language, does not have to satisfy that underlying requirement). Furthermore, when discussing economic viability, the ECJ adopted the test stated by Advocate General Jacobs to the effect that viability

100 The Vertical Restraints Guidelines states that "Dominant companies may not impose non-compete obligations on their buyers unless they can objectively justify such commercial practice within Article 82" (*see* Guidelines on Vertical Restraints, 2000 O.J. (C 291) 1, at para. 141). That statement disregards the necessity to demonstrate a level of coverage that evidences foreclosure. The Discussion Paper helpfully acknowledges that that is an issue to be taken into account and, given that, the same issue should be relevant here. What impact the termination has depends on the facts. While the importance of that customer is the starting point, if the termination is intended or likely to discourage others from dealing with a competitor, then its impact is wider than the sales of that customer alone.

101 Discussion Paper, *supra* note 1, at para. 229.

102 Case C-7/97, Oscar Bronner v. Mediaprint, 1998 E.C.R. I-7791 (ECJ) [hereinafter *Bronner*].

must be assessed on the assumption that the competitor could achieve the same economies of scale as the dominant firm. DG COMP does not make that assumption clear. On the contrary, its test could be read in exactly the opposite sense, namely to rule out that assumption and to allow for a consideration of whether the competitor could in fact achieve comparable economies of scale.

The concept of indispensability, as developed by the ECJ, is designed to strike a balance between the dominant firm's development incentives and rivals' development opportunities. The Court in *Bronner*, and Advocate General Jacobs in particular, stressed the risks to investment if Article 82 is applied too readily to deprive dominant firms of the fruits of their development activities. While DG COMP acknowledges that point (including the need to permit the cost of failed projects to be recovered), it introduces the notion that refusal to supply may only be justified "for a certain period of time in order to ensure an adequate return on . . . investment."[103] That caveat is not to be found in the jurisprudence and is not consistent with the concept of indispensability. That is a function of the competitor's requirement to obtain access to the facility in question, not the dominant firm's requirement to recover its investment. Admittedly, the latter factor led the Court to insist on the requirement of indispensability, but it did not determine the way in which the requirement was expressed.

On the contrary, the concern (articulated by the Advocate General in *Bronner*) that an overly broad rule would be unworkable argues strongly against the introduction of that concept. There is an enduring and unavoidable tension between intellectual property rights and competition law. However well-drawn the boundaries of an intellectual property right may be, it can never ensure that the IP right holder achieves no more than an adequate return in all circumstances. To the contrary, it is a mathematical necessity that it will not be the case because the scope of an IPR represents the legislator's (presumed) best ex ante estimation of the level of protection that it is reasonable to provide across a range of possible outcomes to induce a socially optimal level of innovation. As the range of outcomes includes both stunning success and abject failure, it is unavoidable that some rewards will appear to be excessive. However repugnant the rewards in an individual case may appear to be with the benefit of hindsight, using a competition law remedy to avoid those excesses necessarily carries with it the risk that the legislative balance will be disturbed.

AS THE RANGE OF OUTCOMES INCLUDES BOTH STUNNING SUCCESS AND ABJECT FAILURE, IT IS UNAVOIDABLE THAT SOME REWARDS WILL APPEAR TO BE EXCESSIVE . . . USING A COMPETITION LAW REMEDY TO AVOID THOSE EXCESSES NECESSARILY CARRIES WITH IT THE RISK THAT THE LEGISLATIVE BALANCE WILL BE DISTURBED.

103 Discussion Paper, *supra* note 1, at para. 235.

The dangers of ad hoc and ex post intervention are also apparent in the approach that DG COMP advocates in its assessment of the trade-off between the protection of the original innovator's incentives and the subsequent innovator's opportunities. In its assessment, it makes explicit reference to the need to have regard for the fact that the costs of the original innovation may have been low and the value of the subsequent innovation may be high. No doubt, it can be said that that would only form part of an overall assessment in which the factors on the other side must also be taken into account, but it sends a signal as to the direction in which DG COMP's sympathies may lie.

E. REFUSAL TO LICENSE IPRS: NEW PRODUCT

DG COMP adopts the formula expressed by the ECJ in *IMS Health* that differentiates between those activities that amount to no more than a duplication of the dominant firm's products (that do not justify a compulsory license) and those that involve the development of a new product not offered by the IP right holder and for which there is a potential consumer demand (that may justify a compulsory license).

On this occasion, DG COMP does not add to the language of the Court. That, however, does mean that it does not address the difficult questions about the concept that were not answered by the Court. There is a strong sense that, for both the Court and the Commission, this is a concept that they wish to develop in a reactive way, responding to the different fact patterns that present themselves. Understandable though that may be, it is not wholly consistent with the concept of this Discussion Paper as a source of ex ante guidance to courts, agencies and firms.

F. ALL CASES: OBJECTIVE JUSTIFICATION

The most contentious justification is the dominant firm's desire to integrate downstream, especially where that involves a departure from its established supply arrangements. In that case, DG COMP states that it is the responsibility of the dominant firm to show that termination of the existing supply relationship makes consumers better off than they would be if the existing supply arrangements were to continue either as they are or in competition with the dominant firm.

This issue was canvassed extensively in *Genzyme v. OFT*[104] where the U.K. Competition Appeal Tribunal (CAT) held that the dominant supplier of a drug that required sophisticated home administration and nursing support could not lawfully reintegrate that activity into its own operation. The CAT attached great importance to patients' freedom to obtain such services from the supplier of their choice, even though the previous arrangement had consisted in an exclusive distribution arrangement with a third party.

104 Case 1016/1/1/03, Genzyme v. OFT, 2004 C.A.T. 4.

It appears, therefore, that a dominant firm may only switch to a policy of pure vertical integration if it is the only means by which specified, realized, and provable consumer benefits can be accomplished. For example, it likely would be insufficient to argue:

(a) that to dispense with independent distribution would yield internal efficiencies for the supplier; or

(b) that independent distribution is comparatively less efficient than pure vertical integration

unless, perhaps, it could be shown that those efficiency gains would be passed through to customers in the form of prices lower than those likely to be realized under a competitive or independent distribution system. Given the Commission's assumptions about the effects of dominance, it may be challenging to establish that proposition.

To prevail, it may be necessary for the dominant firm to go further and establish that the existing arrangements have failed in a way that goes beyond considerations of "mere" efficiency (for example, in a way that threatens customer safety) and that the only way in which repetition of that failure can be avoided is through reintegration into the dominant firm.

The logic of this approach is dubious. Analytically, a refusal to supply in the distribution context could be perceived as a tying of the supply and distribution activities that should only merit intervention where one of the three exclusionary effects canvassed by DG COMP in that context is established, requiring a showing of foreclosure in relation to either the supply or distribution market. Except in the linguistic sense that pure vertical integration entails the exclusion of independent distributors from the distribution of that product, none of those effects is likely to be made out. In the special circumstances of a monopoly drug of the kind at issue in *Genzyme*, the exception may be pertinent but, other than in such a case, acceptance of the exception would reintroduce the per se concept that this Discussion Paper eschews.

As is the case throughout this Discussion Paper, the solution to the problem, therefore, lies not in a more extensive interpretation of the concept of objective justification, but in a better analyzed approach to the scope of the abuse.

VII. Conclusion

DG COMP's Discussion Paper is a welcome statement that the control of exclusionary conduct is designed to promote consumer welfare and should be tested by an effects-based analytical model. However, the full value of that step forward has yet to be realized because DG COMP's adoption of a precautionary approach appears to capture every constraint on competitive expansion and limits the scope for justification to the bare minimum. There are, of course, paragraphs and sentences in the Discussion Paper that suggest a less conservative approach. Indeed, it is one of the Discussion Paper's features that it contains something for everyone. The danger is that everyone will look to those parts of the document that suit them and disregard the less palatable parts. It is troublesome that the Discussion Paper does not contain enough to move the enforcement of Article 82 (not simply by the Commission, but also by the national courts and national competition authorities) towards an assessment that segregates those cases where dominant firms' conduct does substantially lessen competition from those where it is in fact a part of the competitive process and does maximize consumer welfare, a part that is essential to accomplishing DG COMP's proclaimed objective given the very fact of dominant firms' market position. Until that is fully accomplished, there can be no assurance that enforcement of Article 82 will remove rather than create constraints upon a dynamic process of competition.

Volume 2 | Number 1 | Spring 2006

Competition Policy International

The Article 82 Review Process and Its Impact on Compulsory Licensing of IP Rights

Frank Montag and Alicia Van Cauwelaert

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

The Article 82 Review Process and Its Impact on Compulsory Licensing of IP Rights

Frank Montag and Alicia Van Cauwelaert

The European Commission is presently reviewing the way in which it regulates the unilateral behavior of companies with market power under Article 82 of the EC Treaty and has published a discussion paper in this regard in December 2005 (Discussion Paper). In line with other areas of EC competition law, it is clear that the Commission is eager to adopt an economics-based approach to Article 82, with the focus being on consumer harm rather than the protection of particular competitors.

This paper reviews the position put forward by the Commission in relation to the concept of an exclusionary abuse, the meaning of dominance, and the use of an efficiency defense. In particular, the paper looks at refusal to supply cases involving IP rights and the impact the Article 82 review may have on such cases in the future. In general, the Discussion Paper does not indicate a change of policy with regard to first-time refusals to supply or license. However, the weight attached to existing commercial arrangements could result in behavior that previously would not have been considered as abusive, falling foul of Article 82. Although not considered in the Discussion Paper, in our view, the "no economic sense test" could be useful in determining whether a refusal to continue supplying an existing customer is objectively justified.

Dr. Frank Montag is a partner and Alicia Van Cauwelaert is an associate in the Antitrust, Competition, and Trade Group of Freshfields Bruckhaus Deringer. Both are based in Brussels. This paper is based on a presentation given by Dr. Frank Montag at the ABA Fall Forum on Nov. 15, 2005. The authors would like to acknowledge their colleague Rachel Brandenburger whose speaking notes from the Post-Fordham International Competition Policy Conference, workshop on Antitrust and Intellectual Property, Oct. 9, 2004 were helpful in drafting the section on compulsory licensing and IP rights. A summary of the workshop on Antitrust and Intellectual Property is available at http://www.iilj.org/events/details/event-intl_competition policy.htm.

I. Introduction

EC competition law has evolved considerably since the provisions on competition law in the Treaty Establishing the European Community (EC Treaty) came into force in 1957. A key theme of this evolution is the increased role played by economic analysis over the last decade that, in recent years, has resulted in a reform of legislation and enforcement practice known under the name of modernization.

Although the text of Article 81 of the EC Treaty, concerning agreements between companies that may restrict competition, has remained unchanged since 1957, the European Commission's enforcement policy and the implementing regulations have undergone considerable change, with a move away from legalistic form-based rules to a more economic effects-based approach. This shift in enforcement policy is demonstrated by the recent Commission guidelines on the effect of trade concept, the guidelines on the application of Article 81(3), as well as the guidelines in relation to horizontal and vertical agreements.[1] This change of the substantive approach was accompanied by procedural reform consisting of the modernization of the implementation legislation that came into force on May 1, 2004.[2]

Similarly, the amendments to the EC merger control regime that entered into force on May 1, 2004, and the accompanying horizontal merger guidelines are proof of a more economics-based approach to merger control.[3] This reform established the new test for the prohibition of mergers (i.e., the test of significant impediment to effective competition, in which the old prohibition criterion of creating or strengthening a dominant position has been downgraded to the function of a mere example of the application of the new test). The new horizontal merger guidelines introduced the concepts of substantial market power as well as unilateral and coordinated effects, thereby bringing the interpretation and appli-

1 Commission Notice on Guidelines on the effect on trade concept contained in Articles 81 and 82 of the Treaty, 2004 O.J. (C101/81) (Apr. 27, 2004); Commission Notice on Guidelines on the application of Article 81 (3) of the Treaty, 2004 O.J. (C101/97) (Apr. 27, 2004); Commission Notice on Guidelines on the applicability of Article 81 of the EC Treaty to horizontal cooperation agreements, 2001 O.J. (C3/02) (Jan. 6, 2001); Commission Notice on Guidelines on Vertical Restraints, 2000 O.J. (C 291/1) (Oct., 10, 2000).

2 Council Regulation (EC) No. 1/2003 of Dec. 16, 2002 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty, 2003 O.J. (L 1/1) (Jan. 4, 2003).

3 Council Regulation (EC) No. 139/2004 of Jan. 20, 2004 on the control of concentrations between undertakings, 2004 O.J. (L 24) (Jan. 29,.2004) and Council Notice on Guidelines on the assessment of horizontal mergers, 2004 O.J. (C 31/5) (Feb. 5, 2004). The Commission is also in the process of amending the rules with regard to the provision of state aid by the EC Member States in order to ensure that aid is permitted or prohibited on the basis of sound economic principles. *See* European Commission, State Aid Action Plan: Less and better targeted state aid: a roadmap for state aid reform 2005 – 2009, *at* http://europa.eu.int/comm/competition/state_aid/others/action_plan/saap_en.pdf (last visited Feb. 8, 2006).

cation of the EC merger regime much closer to current economic thinking and the U.S. practice of merger control.

The importance placed by the Commission on economic analysis and the strengthened role it plays in the application of EC competition rules is also evidenced by the appointment of a Chief Competition Economist and the creation of the Chief Economist team that is currently made up of an additional ten specialized economists all of whom hold Ph.D.s in industrial organization.[4]

In contrast to merger control and the rules on anticompetitive agreements and concerted practices, the law relating to the unilateral behavior of companies with market power remains the one area of EC competition law that has not undergone some degree of modernization in order to reflect this shift in emphasis towards the economic effects of the activities of undertakings. Therefore, it was not surprising when, in 2003, the Commission announced that it would undertake a review of the way in which it regulates the unilateral behavior of companies with market power under Article 82 of the EC Treaty.[5] Since the review of Article 82 was announced, it has become clear that the Commission is eager to adopt an economics-based approach to Article 82 and that a drive for consistency with other areas of competition law is one of the underlying reasons for the review.[6]

The review of Article 82 policy is overdue, given the criticism of cases decided by the Commission and the EC Courts and the considerable amount of uncertainty that exists on the degree of freedom that a company with market power may have and the type of behavior it can lawfully engage in. For example, the conclusion of the Commission and the EC Courts in *Michelin II*[7] and *Virgin/British Airways*[8] that the rebates in question were per se abusive on the basis of their form, and the authorities' failure to undertake an analysis of the actual effects of the behavior in question, does not sit well with the current

4 Speech by Lars-Hendrik Röller, *Using economic analysis to strengthen competition policy enforcement*, Modelling European Mergers: Theory, Competition Policy and Case Studies, European Commission, Sep. 1, 2005, *available at* http://europa.eu.int/comm/competition/speeches/index_2005.html.

5 Speech by Mario Monti (at the time, EC Competition Commissioner), *EU competition policy after May 2004*, Fordham Annual Conference on International Antitrust Law and Policy, New York, Oct. 24, 2003, *available at* http://europa.eu.int/comm/competition/speeches/index_2003.html.

6 Speech by Philip Lowe, *Enforcement Authorities Roundtable*, Fordham Annual Conference on International Antitrust Law and Policy, New York, Sep. 22, 2005.

7 Case T-203/01, Manufacture Française des Pneumatiques Michelin v. Commission of the European Communities, 2003 E.C.R. II-4071 (CFI) [hereinafter *Michelin II*].

8 Case T-219/99, British Airways v. Commission, 2003 E.C.R. II-5917 (CFI) [hereinafter *Virgin/British Airways*].

emphasis on economic effects in the application of Article 81 and EC merger control rules. There is also a difficulty in determining whether certain types of behavior, that have not been specifically considered in previous cases, will be considered abusive or not.

As part of the review of Article 82 policy, the Commission published a discussion paper on the application of Article 82 to exclusionary abuses on Dec. 19, 2005 (Discussion Paper).[9] The Commission has indicated that a review of exploitative and discriminatory abuses will be undertaken in 2006 and so these abuses are not dealt with in the Discussion Paper. The publication of the Discussion Paper follows considerable consultation both within the Commission and with EU national competition authorities as well as other interested bodies and non-EU antitrust enforcers in relation to exclusionary abuses.[10]

The formal publication of the Discussion Paper marks the opening of an even wider consultation process with the Commission inviting comments on the Discussion Paper by Mar. 31, 2006. Current indications from the Commission are that this consultation process could lead to changes to Commission policy on Article 82 in relation to exclusionary practices that are not considered in the Discussion Paper.

There has been much debate as to whether the Commission's review should result in guidelines on the law concerning abuse of dominance or whether it would be preferable that the Commission adopts a more economic approach to its enforcement of Article 82 on a case-by-case basis. In particular, some voices within the Commission's Directorate-General for Competition (DG COMP) and among national competition authorities have pointed out the harm that may be caused by guidelines if the Commission is not able to articulate its policy in a transparent and meaningful manner. The same voices have also noted that harm may also be caused if the guidelines become too detailed and prescriptive, thus forcing the application of Article 82 into a straitjacket—

THERE HAS BEEN MUCH DEBATE AS TO WHETHER THE COMMISSION'S REVIEW SHOULD RESULT IN GUIDELINES ON THE LAW CONCERNING ABUSE OF DOMINANCE OR WHETHER IT WOULD BE PREFERABLE THAT THE COMMISSION ADOPTS A MORE ECONOMIC APPROACH TO ITS ENFORCEMENT OF ARTICLE 82 ON A CASE-BY-CASE BASIS.

9 EUROPEAN COMMISSION, DG COMPETITION DISCUSSION PAPER ON THE APPLICATION OF ARTICLE 82 OF THE TREATY TO EXCLUSIONARY ABUSES (Dec. 2005), *available at* http://europa.eu.int/comm/competition/antitrust/others/discpaper2005.pdf [hereinafter Discussion Paper].

10 *See* Competition Law Forum Article 82 Review Group, *The reform of Article 82: Recommendations on Key Policy Objectives*, 1 EUR. COMPETITION REV. 179 (2005) and *An Economic Approach to Article 82,* Report by the Economic Advisory Group for competition policy commissioned by the Chief Economist of the Competition Directorate of the European Commission (originally published Jul. 2005), reprinted in this issue as J. Gual et al., *An Economic Approach to Article 82,* 2(1) COMPETITION POL'Y INT'L 111-154 (2006) [hereinafter EAGCP Report].

preventing a flexible adjustment of enforcement policy to changing business practices. On the other hand, guidelines on the Commission's policy in relation to Article 82 will not only be of importance to companies in assessing how the Commission will assess certain behavior, but will also provide guidance to national competition authorities and national courts in the 25 EU Member States, who must also apply Article 82, in a consistent manner, to behavior that significantly effects cross border trade within the European Union.[11]

Although the final outcome of the Article 82 review, whether or not it takes the form of guidelines, cannot change EC law as set out in Article 82 or the previous case law of the European Court of Justice (ECJ) or the European Court of First Instance (CFI), it will provide valuable guidance on the way in which the Commission will apply Article 82 in the future and, in particular, its enforcement priorities. However, until it is clear that national competition authorities and national courts will follow the position put forward by the Commission following its review, or until the EC Courts have confirmed that the Commission's approach is correct, there is a risk that although the Commission may be unlikely to take enforcement action under Article 82 against particular behavior it considers acceptable, the behavior could still be found to be in breach of Article 82 by a national court or national competition authority.

Accordingly, until there is further clarity, the publication of the Commission's Discussion Paper does not provide companies with possible market dominance or national courts or national competition authorities with sufficient guidance on the application of Article 82.[12] In fact the publication of the Discussion Paper now places companies with market power, and their advisors, in an even greater quandary as to whether or not to follow the strict approach confirmed by the case law or to examine, based on a more economic analysis, whether the general principles, presumptions, and possible defenses set out in the Discussion Paper indicate that the behavior is acceptable under Article 82.[13]

While there are many aspects of the Article 82 review that give rise to debate, including market definition, the assessment of single or collective dominance, and abusive intent, this paper will focus on the policy objectives behind Article 82, the definition of an exclusionary abuse of a dominant position, and the application of an efficiency defense. The second part of this paper takes a closer look at decisions of the Commission and the EC Courts in relation to a refusal to license an IP right in order to assess if the eventual outcome of the Article 82

11 Article 3(2) of Regulation 1/2003 provides that EC Member States may apply national laws to unilateral conduct which is stricter than Article 82, however, to the extent that there is an effect on cross-border trade, national laws may not permit behavior which is prohibited by Article 82 (*see supra* note 2).

12 This is recognized by the Commission (*see* Discussion Paper, *supra* note 9, at para. 7).

13 In particular, the possibility that not all loyalty rebates offered by a dominant company will be considered abusive, indicates a shift away from the per se approach confirmed in *Michelin II*, *supra* note 7.

review may impact the way such cases are dealt with in the European Community in the future. As evidenced by the divergent views held in relation to the *Microsoft* decision, the distinction between exclusionary abusive behavior and non-abusive behavior is particularly controversial in relation to IP rights where the exclusion of others through the lawful exercise of an IP right can in certain circumstances be deemed unlawful under antitrust law.

II. Article 82 Policy Review

A. POLICY OBJECTIVES

The Commission has repeatedly stated that the objective of Article 82 is "the protection of competition on the market as a means of enhancing consumer welfare and of ensuring an efficient allocation of resources."[14] In principle, this is the same objective that applies to other areas of competition law such as the law on anticompetitive agreements, mergers, and state aid, and should give a welcome coherence to EC competition policy.[15] However, this approach still seems influenced by the traditional application of Article 82 that had the aim to protect the competitive process by preserving a competitive market structure, rather than by focusing directly on consumer harm. The Commission states in the Discussion Paper that "the Commission will adopt an approach which is based on the likely effects on the market."[16] It also explains that:

> "the concern is to prevent exclusionary conduct of the dominant firm which is likely to limit the remaining competitive constraints on the dominant company, including entry of newcomers, so as to avoid that consumers are harmed. This means that it is competition, and not competitors as such that is to be protected."[17]

14 With regard to exclusionary abuses, see Discussion Paper, *supra* note 9, at paras. 4 and 54; Philip Lowe, *supra* note 6; Speech by Neelie Kroes, *European Competition Policy – Delivering Better Markets and Better Choices*, European Consumer and Competition Day, London, Sep. 15, 2005; and Speech by Neelie Kroes, *Preliminary Thoughts on Policy Review of Article 82*, Fordham Corporate Law Institute, New York, Sep. 23, 2005, *available at* http://europa.eu.int/comm/competition/speeches/index_2005.html.

15 Philip Lowe, *supra* note 6.

16 Discussion Paper, *supra* note 9, at para. 4.

17 *Id.* at para. 54.

The above formulation of the objective and the effects-based application of Article 82, in practice, should mean that authorities only intervene under EC competition law to protect competition where a failure to do so would result in consumer harm. However, the ultimate outcome of the Article 82 review needs to further clarify the extent to which the focus will be on harm to consumers, rather than harm to the competitive process in the absence of consumer harm. The extent to which the Commission intends to focus on consumer harm, or to presume consumer harm where there are no actual or likely anticompetitive effects on the market, is not made clear in the Discussion Paper. In this regard, the Discussion Paper provides that "harm to intermediate buyers is generally presumed to create harm to final consumers."[18] EC Competition Commissioner Kroes, on the other hand, when announcing the preliminary results of the Article 82 review in a 2005 speech at Fordham Corporate Law Institute, stated that "ultimately the aim is to avoid consumer harm."[19] Commissioner Kroes went on to stress further the position of consumers by adding "I like aggressive competition – including by dominant companies – and I don't care if it may hurt competitors – as long as it ultimately benefits consumers." Commissioner Kroes, however, has indicated that it will be sufficient if there is "likely" to be harm to consumers in the medium or long term indicating that EC officials still intend to take a longer-term approach to consumer harm than their U.S. counterparts. Similarly, the Commission in the Discussion Paper refers to harm to consumers in a "direct or indirect way" and that "not only short term harm, but also medium and long term harm arising from foreclosure is taken into account."[20]

It is possible that the emphasis on the prevention of consumer harm that is "likely" to occur in the future may indicate that with regard to unilateral behavior of a company with market power, the Commission continues to focus on avoiding decisions which wrongly permit anticompetitive behavior (known as type I errors or false negatives).[21] On the other hand, current economic thinking seems to suggest that, with respect to the regulation of the behavior of companies with substantial market power, the emphasis should be on the need to avoid decisions which wrongly condemn pro-competitive behavior (known as type II errors or false positives). However, the distinction between avoidance of type I or type II errors may ultimately be superfluous if, as pointed out in a report prepared by the Economic Advisory Group for Competition Policy (EAGCP Report), an economics effects-based approach, correctly applied, reduces the

18 *Id.* at para. 55.

19 Neelie Kroes (Sep. 23, 2005), *supra* note 14.

20 Discussion Paper, *supra* note 9, at para. 55.

21 Although it is recognized that it is not an offence to hold a dominant position, the tone of the Discussion Paper still indicates a certain mistrust of dominant companies.

likelihood of both condemning pro-competitive behavior and permitting anti-competitive behavior.[22]

Traditionally, EC competition law has also been driven by other goals such as the achievement of the internal market within the European Community, the protection of small and medium sized enterprises, fairness, and successful market liberalization through the privatization of state run industries.[23] The Discussion Paper leaves the door open for these objectives to continue to play a role in EC competition law by explaining that, to give an example, the achievement of market integration will enhance consumer welfare "since the creation and preservation of an open single market promotes an efficient allocation of resources throughout the Community for the benefit of consumers."[24] The extent to which these objectives should play a role in Article 82 policy should be looked at closely and, if they are to continue to play a role, the manner in which they will influence enforcement policy, should be set out clearly and transparently, without risk of being perceived as a hidden agenda.[25]

B. DOMINANCE

The classic definition of dominance in the case law of the EC courts[26] is that an undertaking enjoys a position of economic strength "which enables it to hinder the maintenance of effective competition on the relevant market by allowing it to behave to an appreciable extent independently of its competitors and customers and ultimately of consumers." In the Discussion Paper, the Commission uses this definition, and interprets the notion of independence contained in the definition to mean that the dominant undertaking must have substantial market power.[27] Similar to its approach in the horizontal merger guidelines, the Commission defines substantial market power as the power to influence prices,

---

22 *See supra* note 10.

23 Former EC Competition Commissioner, Mario Monti has stated that:

> The liberalisation process that the Community has launched in recent years can only be successfully achieved if former monopolists, who usually retain powerful market positions, are prevented from engaging in exclusionary practices that delay or prevent effective competition in these markets

(Mario Monti, *supra* note 5).

24 Discussion Paper, *supra* note 9, at para. 4.

25 *See* Brian Sher, *The Last of the Steam-Powered Trains: Modernising Article 82,* 25(5) EUR. COMPETITION L. REV. 263-285 (2004).

26 Case 27/76, United Brands v. Commission, 1978 E.C.R. 207; Case 85/76, Hoffmann-La Roche v. Commission, 1979 E.C.R. 461 [hereinafter *Hoffman-La Roche*]; Case 322/81, Michelin v. Commission, 1983 E.C.R. 3461, at para. 30.

27 Discussion Paper, *supra* note 9, at para. 23.

output, innovation, the variety or quality of goods and services, or other parameters of competition in the market for a significant period of time.[28]

Despite the acknowledgment in the Discussion Paper that in most cases the dominance analysis needs to be extended beyond market shares, encompassing an analysis of competitors, barriers to entry and expansion, and the market power of buyers, there is a concern that the Discussion Paper continues to rely heavily on a presumption of dominance based on market shares.[29] The Discussion Paper provides that "it is very likely that very high market shares, which have been held for some time, indicate a dominant position."[30] This would be the case where a firm holds 50 percent or more of a market, but could also apply in the range of 40 percent to 50 percent. The Commission also indicates that undertakings with a market share below 40 percent[31] also may be considered dominant, although dominance is not likely below 25 percent.

The Discussion Paper does not appear to take into account Commissioner Kroes' remark at Fordham that high market shares are not on their own sufficient to conclude that a dominant position exists, and that a pure market focus risks failing to take proper account of the degree to which competitors can constrain the behavior of the allegedly dominant firm. Indeed economically, the most important factors for the determination of dominance or substantial market power are the existence or absence of barriers to expansion or entry. Even a firm with market shares well above the 50 percent level may not be able to charge supra-competitive prices if it is in constant fear of market entry or capacity expansion by its rivals. That being said, market share thresholds can play a useful role when they are used to define safe havens for firms that would allow them to determine, without a full economic analysis, that they are not subject to the special rules of Article 82.

C. CONCEPT OF EXCLUSIONARY ABUSE

Article 82 prohibits exclusionary, exploitative, and discriminatory abuses of a dominant position. As explained by Commissioner Kroes, the Commission has given priority to the review of exclusionary abuses on the basis that exclusion is often at the basis of later exploitation of customers.[32] Exploitative and discrimi-

28 *Id.* at para. 24.

29 *Id.* at para. 31.

30 *Id*.

31 *See id.* at para. 31 (citing Case C-250/92, Gøttrup-Klim v. DLG, 1994 E.C.R. I-5641, at para. 48, where the undertaking concerned held shares of 36 percent and 32 percent, and the ECJ stated that an undertaking holding market shares of that size may, depending on the strength and number of its competitors, be considered to be in a dominant position).

32 Neelie Kroes (Sep. 23, 2005), *supra* note 14.

natory abuses shall be looked at in the second round of the Article 82 policy review.[33] Another reason for the focus on exclusionary abuses could be the fact that the great majority of past decisions by the Commission and the EC Courts have concerned exclusionary as opposed to exploitative or purely discriminatory practices.

With regard to defining what is an exclusionary abuse, the Commission in the Discussion Paper continues to use the definition provided by the EC Court in *Hoffmann-La Roche*, namely:

> "abuse is an objective concept relating to the behavior of an undertaking in a dominant position which is such as to influence the structure of a market . . . and which, through recourse to methods different from those which condition normal competition in products or services on the basis of the transactions of commercial operators, has the effect of hindering the maintenance of the degree of competition still existing in the market or the growth of that competition."[34]

This has been interpreted by the Commission to mean:

(1) the conduct must be capable of foreclosing rivals; and

(2) in the specific market context, gives rise to a likely market distorting foreclosure effect.[35]

The Discussion Paper goes on to explain that by market foreclosure it means that actual or potential competitors of the allegedly dominant firm are completely or partially denied profitable access to a market. In a move away from a form-based approach, the Commission states in the Discussion Paper that to establish foreclosure it, in general, would be necessary not only to consider the nature of the form of the conduct, but also its incidence, the degree of dominance, and other market characteristics including the existence of network effects and economies of scale.[36] Commissioner Kroes in her speech at Fordham also indicated that foreclosure of one or two competitors would not give rise to a foreclosure effect where sufficient residual competition remained.

33 *Id.*

34 *Hoffman-La Roche*, *supra* note 26.

35 Neelie Kroes (Sep. 23, 2005), *supra* note 14 and Discussion Paper, *supra* note 9, at para. 58 et seq.

36 Discussion Paper, *supra* note 9, at para. 59

The Commission has included a general presumption in the Discussion Paper that provides that where "conduct is clearly not competition on the merits, in particular conduct which clearly creates no efficiencies and which only raises obstacles to residual competition, such conduct is presumed to be an abuse."[37] However, one wonders whether this presumption adds any greater clarity than there is today about the type of behavior that is prohibited by Article 82—in particular given the uncertainty and lack of clarity concerning the meaning of the phrase "competition on the merits."

With regard to pricing behavior, the Commission indicates in the Discussion Paper that the use of an "as efficient competitor test" may be helpful in establishing whether competition is "on the merits."[38] There is a presumption that if a hypothetical competitor who is as efficient as the dominant company could compete against the price schedule or rebate system of the dominant company, the Commission normally will conclude that the pricing behavior constitutes competition on the merits and is not abusive. The "as efficient" competitor test, therefore, creates a kind of "safe harbor" for dominant companies in assessing the level of rebates that they are permitted to offer. On the other hand, if a hypothetical "as efficient" competitor could not compete as a result of the rebates offered by the dominant company, a closer examination of the impact of the behavior will be undertaken.

The Discussion Paper provides guidance on how, where cost information is available, the "as efficient" test is to be applied.[39] However, the difficulty in obtaining the necessary economic evidence on costs in some industries is recognized in the Discussion Paper as well as the fact that in some industries, economies of scale or a "first mover advantage" (particularly in newly liberalized industries) need to be taken into account and competitors cannot be expected to be "as efficient" as the incumbent operator, at least in the short run. Given the exceptions to the "as efficient" test and the need to apply the test in its specific market context, its value as a rule of thumb in determining that certain pricing behavior is acceptable may be limited. Other commentators, while recognizing the value of the "as efficient competitor test" have also highlighted the ambiguous consumer welfare effects of the test, where it can sometimes be in consumers' interests to ensure there is vigorous competition between two firms, even if one of them is less efficient, than to allow the emergence of one monopolist.[40]

---

37 *Id.* at para. 60.

38 *Id.* at para. 63.

39 *Id.* at paras. 64-68.

40 Speech by J. Vickers, *Abuse of Market Power,* European Association for Research in Industrial Economics, Berlin, Sep. 3, 2004, *available at* http://www.oft.gov.uk/NR/rdonlyres/948B9FAF-B83C-49F5-B0FA-B25214DE6199/0/spe0304.pdf and A. Fletcher, *The reform of Article 82: recommendations on key policy objectives,* Competition Law Forum, Mar. 15, 2005, *available at* http://www.oft.gov.uk/NR/rdonlyres/41543043-1363-4CD5-B21E-87A9A3DFD160/0/spe0305.pdf.

It is not yet clear if other tests that have been discussed in the literature will play a role in the enforcement of Article 82.[41] For example, the "profit sacrifice" test (also referred to as the "no economic sense" or "but-for" test) looks at whether the behavior of the dominant company would be profitable or make economic sense in the absence of its tendency to eliminate or lessen competition. The value of the profit sacrifice test with regard to the identification of abusive behavior has been discounted by some commentators as it has been noted that not all exclusionary conduct involves a sacrifice of profit and the test does not help in determining which behavior would not make economic sense in the absence of the foreclosure effect.[42] The EAGCP Report indicates, however, that with regard to certain abuses in the same market, the profit sacrifice test may be useful. In particular, this test may be useful in determining the intent of the company, although it should be noted that abusive intent is not a requirement under EC law. That being said, intent can be taken into account to strengthen an abuse finding,[43] and perhaps as put forward by Amelia Fletcher (2005), lack of intent could, in the absence of evidence of market foreclosure, indicate that the behavior is not abusive.[44] The possibility of using the profit sacrifice test (or more aptly named "no economic sense" test) to determine if a refusal to supply is objectively justified is raised in the second part of this paper.

Another candidate for a standard test is the "consumer harm test" or "consumer welfare test." While this test would be in line with the objectives of EC competition policy it does not assist in identifying the behavior that may lead to consumer harm and leads to further questions as to the standard of proof and whether such harm needs to be actual or potential, or likely or possible.[45]

It has also been suggested that behavior only should be found to be an exclusionary abuse where the conditions set out in Article 82(b) have been satisfied, namely that the behavior of the dominant company "[limits] production, markets or technical development to the prejudice of consumers."[46] The limitation in production can refer to either its own production or that of third parties. The

41 *See id.* and J. Temple Lang & R. O'Donoghue, *The Concept of an Exclusionary Abuse under Article 82 EC,* GCLC Research Papers on Article 82 EC (Jul. 2005).

42 Vickers, *supra* note 40.

43 In the past, the Commission and the EC Courts have looked at intent to support a finding of abuse (*see* Case C-62/86, Akzo v. Commission, 1991 E.C.R. I-3359).

44 Fletcher, *supra* note 40.

45 Vickers, *supra* note 40.

46 This test was first proposed by John Temple Lang. *See* John Temple Lang, *Anticompetitive Non-Pricing Abuses Under European and National Antitrust Law*, *in* International Antitrust Law & Policy (B. Hawk ed., Fordham Corporate Law Institute, 2003), at 235. The test is developed further in Temple Lang & O'Donoghue, *supra* note 41.

requirement that the behavior be detrimental to consumers would appear to be in line with the emphasis in Commissioner Kroes' speech at Fordham on the need for consumer harm. However, the proposition that a dominant company only should be prohibited from creating an obstacle or handicap that would not otherwise exist needs to be debated further.

It would be most welcome if a standard test could be developed in order to indicate with certainty whether behavior was exclusionary and accordingly prohibited under Article 82 while at the same time taking into account the economic effects of the behavior in question. However, such a "one size fits all test" is likely to give rise to either too many type I or type II errors and so an appropriate balance must be sought between a practical transparent test that may give rise to errors and one that, although economically sound in principle, leads to uncertainty and a danger of an ad hoc approach to each case by competition authorities and national courts. There does not appear to be any clear consensus yet on the use of a standard test to determine what behavior is exclusionary. Although one standard test would be welcomed, in practice, it may be necessary to look at different tests for different types of exclusionary conduct.

There does not appear to be any clear consensus yet on the use of a standard test to determine what behavior is exclusionary. Although one standard test would be welcomed, in practice, it may be necessary to look at different tests for different types of exclusionary conduct.

The application of any test will also need to take into account the appropriate standard of proof. For a number of years there has been much speculation in the European Community about the applicable standard of proof in Article 82 cases. Recent case law of the EC Courts[47] has indicated a rather low standard of proof referring to behavior as abusive where it was "capable of" having exclusionary effects or that "tends to" have exclusionary effects and has been much criticized. Commissioner Kroes on the other hand has referred to the need to show "actual or likely" restrictive effects. Similarly, the Discussion Paper refers to "actual or likely anticompetitive effects" and the fact that the Commission approach will be based on the "likely effects on the market."[48]

Although it is not dealt with in the Discussion Paper or in Commissioner Kroes' speech, it is arguable that the economic evidence that the Commission will need to rely on in Article 82 cases will need to meet the standard of proof set out by the Court in relation to decisions made under the EC merger control rules. The Court recently clarified that where the Commission wishes to prohibit a merger under the EC merger control rules, the Commission's evidence must

47 *Michelin II, supra* note 7 and *Virgin/British Airways, supra* note 8.

48 Discussion Paper, *supra* note 9, at paras. 4 and 55.

be "factually accurate, reliable and consistent," contain all the information necessary to assess a complex situation, and be capable of substantiating the Commission's conclusions.[49] The Court added:

> "42. A prospective analysis of the kind necessary in merger control must be carried out with great care since it does not entail the examination of past events – for which often many items of evidence are available which make it possible to understand the causes – or of current events, but rather a prediction of events which are more or less likely to occur in future if a decision prohibiting the planned concentration or laying down the conditions for it is not adopted.
> 43. Thus, the prospective analysis consists of an examination of how a concentration might alter the factors determining the state of competition on a given market in order to establish whether it would give rise to a serious impediment to effective competition. Such an analysis makes it necessary to envisage various chains of cause and effect with a view to ascertaining which of them are the most likely."

Where the Commission seeks to apply Article 82 to conduct that may have more or less likely harmful effects in the future, it will have to satisfy the standards established by the Court in *Tetra Laval*. There is no reason to assume that the standard for proving likely future effects in abuse of dominance cases should be any lower than in merger cases.

D. EFFICIENCY DEFENSE

The possibility of weighing the anticompetitive and the pro-competitive effects against each other in an Article 82 analysis has been the subject of much discussion and debate in the past. Previously, Commission officials and other commentators have noted that the text of Article 82 makes no provision for allowing abusive behavior—even if it is in the long-term interest of consumers.[50] However, Commissioner Kroes states, "we must find a way to include efficiencies in our analysis."[51] In the past, it has been open to dominant companies to explain that their behavior was "objectively justified" and accordingly not abusive. While it may be possible to bring an efficiency defense within this objective jus-

49 Case C-12/03P, Commission v. Tetra Laval (Feb. 15, 2005, not yet reported) [hereinafter *Tetra Laval*].

50 Speech by Philip Lowe, Fordham Annual Conference on International Antitrust Law and Policy, New York, Oct.23, 2003, *available at* http://europa.eu.int/comm/competition/speeches/index_2003.html.

51 Neelie Kroes (Sep. 23, 2005), *supra* note 14.

tification test, some commentators have noted that this may be inadequate as the objective justification defense does not, on its face, allow a weighing up of the benefits and the anticompetitive effects.[52] Similarly, the Discussion Paper distinguishes between an efficiency defense and the two objective justification defenses (i.e. the so-called "objective necessity defense" and the so-called "meeting the competition defense").[53]

The "efficiency defense," outlined in the Discussion Paper, does not differ from the framework set out in relation to Article 81 on restrictive agreements and would require a dominant company whose behavior is being examined to demonstrate that conditions similar to those attached to Article 81(3) are satisfied. In summary, the particular behavior that is potentially abusive must meet the following conditions:

(1) the conduct must give rise to specific efficiencies;

(2) the conduct must be indispensable to the attainment of those efficiencies;

(3) the benefits must outweigh the negative effects;

(4) the benefits must be passed on to consumers (or at the very least consumers must not be worse off); and

(5) all competition must not be eliminated.

The requirement that all competition must not be eliminated in order to satisfy an efficiency defense has lead Commissioner Kroes to indicate that there might be a level of super dominance at which the efficiency defense will never be successful. The Discussion Paper confirms that it is "highly unlikely that abusive conduct of a dominant company with a market position approaching that of a monopoly, or with a similar level of market power could be justified on the ground of efficiency gains."[54] It goes on to say that a company is considered to have a market position approaching that of a monopoly if its market share exceeds 75 percent and if there is almost no competition left from actual competitors in the market. Commissioner Kroes also notes that there are some types of abusive behavior for which there are no efficiencies at all, such as misuse of the patent system and the provision of misleading information to patent authorities as described in the recent Commission decision against AstraZeneca.[55]

---

52 Fletcher, *supra* note 40. However, Fletcher also notes that the discussion of objective justification was wider in the *Microsoft* decision which is discussed further in Section III of this paper.

53 Discussion Paper, *supra* note 9, at para. 78 et seq.

54 *Id.* at para. 91.

55 Press Release, European Commission, IP/05/737 Competition: Commission fines AstraZeneca £60 million for misusing patent system to delay market entry of competing generic drugs (Jun. 15, 2005) and Neelie Kroes (Sep. 23, 2005), *supra* note 14.

It appears that the Commission envisages looking only at efficiencies raised by the company under investigation as a defense once an abuse has been established. Among other things, this has an effect on the burden of proof which is shifted from the Commission to the allegedly dominant company. However, it would seem to be more economically sound to take efficiencies into account when determining whether behavior is abusive. In our view, the efficiency analysis is most appropriately carried out as part of the determination of consumer harm. If certain conduct leads to significant efficiencies, it is unlikely that the Commission would find consumer harm and, therefore, this behavior should not be deemed abusive despite its potential foreclosure effects. Carrying out the efficiency analysis as part of the abuse analysis also would be more legally sound given that the possibility of taking efficiencies into account is absent from the text of Article 82.[56] An efficiency analysis, as opposed to an efficiency defense, may cause the Commission to take efficiencies into account at an early stage of its analysis rather than waiting until all the other elements of abuse have been established.

III. Compulsory Licensing of IP Rights

A. OVERVIEW

The interface between the exclusivity granted by IP rights and the obligation under competition law of the holder of an IP right to license it to third parties has been the subject of much debate following the Magill and IMS Health judgments of the ECJ and most recently the Microsoft decision of the Commission.[57] Because of the many remaining questions in this area, the business community and practitioners of competition law would welcome a clarification of the rules in the framework of the Article 82 review. In order to consider if the Article 82 review may bring about such a clarification of the rules with regard to the compulsory licensing of IP rights in the European Community, it is necessary to first look at the circumstances in which the Commission or EC Courts have considered compulsory licensing.

Traditionally, competition regulators have been reluctant to order the compulsory licensing of IP rights as the IP rights owner's freedom to refuse to grant a

56 With regard to objective justifications, in *GlaxoSmithKline*, Advocate General Jacobs, while noting that Article 82 does not contain any explicit provision for the exemption of conduct otherwise falling within it, stated that in his view it was more accurate to say that certain types of conduct on the part of a dominant undertaking do not fall within the category of abuse at all as a result of its objective justification. This would appear to apply equally to efficiencies (Case C-53/03, Syfait v. GlaxoSmithKline, 2005 E.C.R I-4609 [hereinafter *GlaxoSmithKline*], at para. 72).

57 Joined Cases C-241/91P and C-242/91P, RTE and ITP v. Commission, 1995 E.C.R. I-743 [hereinafter *Magill*]; Case C-418/01, IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. KG, 2004 E.C.R. I-05039 [hereinafter *IMS Health*]; Commission Decision, Case COMP/C-3/37.792, Microsoft (Mar. 24, 2004, not yet reported) [hereinafter *Microsoft*].

license is at the very heart of an IP right. Competition regulators seek to strike a balance between ensuring that antitrust policy does not stifle or chill innovation by reducing the value of an IP right through compulsory licensing, and ensuring that a holder of IP rights does not, through anticompetitive behavior, prevent competition in the marketplace (which will also stifle innovation in the long run). As indicated in the Discussion Paper, enforcement policy towards refusals to supply has to take into account both the effect of having more short-run competition and the possible long-run effects on investment incentives.[58] In light of the overall prominence given to consumer welfare, the Discussion Paper also makes it clear that for a refusal to supply to be abusive, it has to have a likely anticompetitive effect on the market which is detrimental to consumer welfare.[59]

Generally, the refusal to license an IP right is regarded as a subset of the so-called "refusal to supply" or "obligation to deal" category of cases.[60] The obligation to deal has arisen on the basis of Article 82(b) of the EC Treaty that states that it is an abuse for a dominant company to "[limit] production, markets or technical development to the prejudice of consumers." In relation to IP rights, the ECJ in *Volvo v. Veng*[61] held that a refusal to supply a license allowing third parties to manufacture spare parts was not an abuse, in the absence of other abusive or exclusionary conduct, such as an arbitrary refusal to supply spare parts to independent repairers or the fixing of prices at an unfair level.[62] However, in *Magill* and *IMS Health*, the Court found that a refusal to license can be prohibited, even in the absence of other exclusionary conduct, in certain "exceptional circumstances."[63] Although it has been questioned whether the low quality of the IP rights—which are the subject matter of these decisions—had an impact on their outcome, they do set the framework within which cases involving a refusal to license an IP right must be examined. The impact that the nature of the IP rights can have on cases is also looked in more detail later in this paper.

The Commission decision requiring *Microsoft* to inter alia provide interoperability information necessary for competitors to be able to compete effectively in the workgroup server operating system market has been very controversial on

58 Discussion Paper, *supra* note 9, at para. 213.

59 *Id.* at para. 210.

60 *Id.* at para. 209.

61 Case 238/87, AB Volvo v. Erik Veng (U.K.) Ltd., 1988 E.C.R. 6211 [hereinafter *Volvo v. Veng*].

62 The approach taken in Europe in *Volvo v. Veng* is perhaps closest to the position in U.S. law. It appears that is was the absence of a separate abuse by Verizon under the Sherman Act which lead the Court in *Trinko* to conclude that the refusal to deal did not violate antitrust laws Verizon Communications Inc., Petitioner v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398 (2004).

63 *See supra* note 57.

both sides of the Atlantic and is currently under appeal in the EC Courts.[64] It is not clear yet which category of case this part of the *Microsoft* decision falls into. It may concern a simple refusal to supply information, a refusal to license IP rights, or exclusionary conduct that seeks to leverage dominance in one market into another market where the remedy is a compulsory license. This paper looks at the *Microsoft* decision as if it has been accepted that it concerns a refusal to license an IP right. Despite the controversy surrounding the *Microsoft* decision, there appears to be considerable overlap between the Commission's interoperability remedy and the U.S. consent decree.[65]

B. EXCEPTIONAL CIRCUMSTANCES WARRANTING A COMPULSORY LICENSE

The ECJ has found that in "exceptional circumstances" it can be an abuse of dominance for a dominant company to refuse to deal because of the resulting limitation on production, markets, or technical development. The Court in *Magill* held that there were "exceptional circumstances" justifying a compulsory license where certain conditions were satisfied. These conditions were confirmed by the Court in *IMS Health* where it clarified that a refusal to give access to a product or service protected by an IP right can give rise to an abuse where the following cumulative conditions are satisfied:

(1) access to the IP right is indispensable to carrying out that business;

(2) the refusal to license the IP right is preventing the emergence of a new product for which there is a potential consumer demand;

(3) the refusal cannot be objectively justified; and

(4) the refusal would eliminate all competition in a secondary market.

On the other hand, the Commission decision in *Microsoft*, which was made just one month before the *IMS Health* judgment, found that the conditions in *Magill* were not exhaustive and that other conditions, such as the disruption of previous levels of supply, also could be relevant. Although the *Microsoft* decision leaves open the extent to which Microsoft's interface information contains information covered by IP rights, Commission commentators have said that the *Microsoft* decision, in any event, meets the *Magill*/*IMS Health* requirement of exceptional circumstances.[66]

---

64 *Microsoft*, *supra* note 57 and Pending Case T-201/04, Microsoft v. Commission.

65 *See* Press Release, U.S. Department of Justice, Assistant Attorney General for Antitrust, R. Hewitt Pate, Issues Statement on the EC's Decision in its Microsoft Investigation (Mar. 24, 2004), *available at* http://www.usdoj.gov/atr/public/press_releases/2004/202976.htm.

66 Speech by Jürgen Mensching, *The Microsoft Decision – Promoting Innovation*, 4th Annual Competition Law Review Conference, Oct. 22, 2004, *available at* http://europa.eu.int/comm/competition/speeches/index_2004.html.

The Discussion Paper divides refusal to deal cases into four categories, namely, refusal to continue an existing supply relationship; refusal to start supplying; refusal to license an IP right; and refusal to supply information for interoperability. Although the section in the Discussion Paper on refusal to license an IP right formulates the *IMS Health* conditions slightly differently, the scope is effectively the same.[67] However, it is not clear on the face of the Discussion Paper whether a refusal to continue an IP license would be considered as a refusal to continue an existing supply relationship or as a refusal to license an IP right.

The Commission makes it clear in the Discussion Paper that it considers it much easier to show that a termination of an existing relationship is abusive as opposed to a refusal to license an IP right. The Discussion Paper provides that the termination of an existing relationship by a dominant company will be abusive where only the following two conditions are met:

(1) the refusal is likely to have a negative effect on competition; and

(2) the refusal is not justified objectively or by efficiencies.[68]

If this category also covers existing IP licenses, then it would not be necessary, as set out in *IMS Health* and the Discussion Paper in relation to a first-time license of an IP right, to show:

(1) that the supply is indispensable to normal economic activity in the downstream market; and

(2) that the termination will prevent the development of the market to the detriment of consumers (although this would normally mean the prevention of the production of a new product, it could also refer to the prevention of the continued production of a product).

If this is the case, it would appear that the Commission considers it considerably easier to prove an abuse under Article 82 which involves the continuation of an existing IP license as opposed to the first-time license of an IP right.

Existing court jurisprudence does not provide direct guidance on the termination of an existing IP right and whether or not this should be treated in a similar manner to a first-time license. However, the reasons given in the Discussion

67 According to the Discussion Paper, five conditions have to be fulfilled in order for a refusal to start to supply to be abusive: (i) the behavior can be properly characterized as a refusal to supply; (ii) the refusing undertaking is dominant; (iii) the input is indispensable; (iv) the refusal is likely to have negative effects on competition; and (v) the refusal is not objectively justified. In the case of the refusal to license an IP right, an additional, sixth condition must be fulfilled, namely that the refusal to license the IP right prevents the development of a market for which the license is an indispensable input (in other words, the refusal prevents the development of a new product for which there is consumer demand). *See* Discussion Paper, *supra* note 9, at paras. 224 and 239.

68 *Id.* at para. 218.

Paper and the case law for treating IP rights with caution and only ordering compulsory licensing in exceptional circumstances would appear to apply equally to existing licenses and first-time licenses.[69] However, the fact that a license has been granted previously may make it more difficult to objectively justify a refusal to continue the license.

C. NEW PRODUCT

In essence, the limitation on production or technical development, that is considered an abuse under Article 82(b), arises in the case of IP rights (or arguably at the very least in the case of the license of an IP right that had not previously been licensed to the requesting party) where the refusal results in the prevention of the emergence of a "new product" for which there is potential customer demand. It was the prevention of the emergence of a "new product" in the market in *Magill* which distinguished that case from *Volvo v. Veng*, where a refusal to license was only an abuse in combination with other abusive conduct.[70]

The requirement that the IP license is indispensable for the creation of a new product means that the incidences in which EC competition law regulators will require compulsory licensing will be very limited (or at the very least will be very limited in the absence of an existing commercial relationship). However, the extent to which regulators will resort to compulsory licensing will depend to a large degree on how the concept of "new product" is defined. Many had hoped that the Court in *IMS Health* would clarify what is meant by a "new product," but the guidance provided by the Court is limited to the requirement that the company requesting the license must "not intend to limit itself essentially to duplicating the goods or services already offered."[71] However, it is not clear from the case law whether very slight improvements to a product will constitute a new product or whether the new product has to be so different that it actually would compete in a new product market (or indeed, as may be more likely, is somewhere in between).[72]

The Discussion Paper interprets the "new product" requirement to mean that the refusal must not prevent the development of the market for which the license is an indispensable input, to the detriment of consumers. Although the Discussion states that this may only be the case, as indicated in *IMS Health*—where the undertaking requesting the license does not intend to limit itself to

69 *Id.* at para. 238 and *Volvo v. Veng*, *supra* note 61.

70 *See* Opinion of the Advocate General, *IMS Health*, *supra* note 57.

71 *Id.* at para. 49.

72 Damien Geradin, *Limiting the scope of Article 82 EC: What can the EU learn from the U.S. Supreme Court's judgment in Trinko in the wake of Microsoft, IMS and Deutsche Telekom?*, 41 COMMON MARKET L. REV. 1481-1518 (2004).

duplicating the goods or services already offered by the dominant company, but intends to produce new goods or services not offered by the owner of the right for which there is a potential consumer demand[73]—arguably it also could cover a situation that would result in the prevention of the continued production of a product for which consumer demand would otherwise not be met. The Discussion Paper also indicates that it may be abusive to refuse to license an IP right where that IP right is indispensable for follow-on innovation.[74] However, the Discussion Paper appears to go further than existing case law by indicating that it is not necessary for the requesting party to have already identified the potential new product that it wishes to develop using the IP right. This could potentially increase the number of incidences an IP rights holder may be required to license an IP right, as in a number of areas it may be difficult to show that the IP right is not indispensable for the development of a product that has not yet been identified.

THE REQUIREMENT THAT THE IP RIGHT IS, AMONG OTHER THINGS, INDISPENSABLE TO THE CREATION OF A NEW PRODUCT ALSO DISTINGUISHES IP RIGHTS FROM OTHER TYPES OF PROPERTY WHERE THERE MAY BE AN OBLIGATION TO DEAL.

The requirement that the IP right is, among other things, indispensable to the creation of a new product also distinguishes IP rights from other types of property where there may be an obligation to deal. In *Bronner*,[75] the Court looked at the obligation to provide access to a distribution system and found that, other than the requirement that access be required to create a new product, the same conditions that applied in *Magill* needed to be satisfied before a refusal to deal on its own, in relation to physical property, could give rise to an obligation to deal.

The Commission did not deal specifically with the creation of a "new product" in the *Microsoft* decision. However, in its analysis of whether or not the refusal was objectively justified, the Commission did refer to the fact that Microsoft's refusal to supply has resulted, and will continue to result, in blocking new functions of operating systems.[76] It is now up to the Court to decide whether or not the creation of new functions constitutes a "new product" under the conditions laid down in *IMS Health*, or whether the basis that Microsoft had refused to continue to supply information that it had supplied in the past means that the

73 Discussion Paper, *supra* note 9, at para. 239.

74 *Id.* at para. 240.

75 Case C-7/97, Oscar Bronner v. Mediaprint, 1998 E.C.R. I-7791 (ECJ) [hereinafter *Bronner*].

76 The Director within the Commission responsible for the decision has stated that the disclosure of the information in question would allow the development of new products. *See* Jürgen Mensching, *supra* note 66.

Commission is not required to show that the refusal resulted in the prevention of the creation of a new product.

D. OBJECTIVE JUSTIFICATION FOR THE REFUSAL: AN EFFICIENCY DEFENSE?

A dominant company may refuse to license an IP right or supply a good where the refusal can be objectively justified. For example, there is no obligation to supply companies that have a bad payment track record or where supply may damage reputation or goodwill. Similarly, capacity constraints may provide an objective justification for a refusal. The Discussion Paper does not indicate any change in the Commission's position that an objective justification can mean that behavior which would otherwise be considered an exclusionary abuse can be permitted.

Where there has been a previous history of dealing it may be harder for the dominant company to explain why there has been a change of circumstances that objectively justify a refusal to continue to supply. The emphasis put by the Commission in the *Microsoft* decision, and in the Discussion Paper, on the fact that there was a "disruption of previous level of supplies" may lead a dominant company to be more reluctant to license IP rights if it thinks that, in the future, it may wish to refuse to continue licensing the IP rights for its own (perhaps only subjectively justified) reasons.[77] Similarly, the U.S. courts regard a refusal to continue to deal as more objectionable than a refusal to commence dealing.[78] A refusal to continue to deal may make no economic sense but for the exclusionary effects. It is perhaps worth considering the value of the "no economic sense" test in determining whether a refusal to supply, or to continue to supply, can be objectively justified. If the refusal to supply makes economic sense even in the absence of exclusionary effects, it may indicate that the refusal is objectively justified. However, in the Discussion Paper, the Commission suggests that where a dominant company argues that it is terminating a supply relationship because it wants to integrate downstream, it must "show that consumers are better off with the supply relationship terminated."[79] Accordingly, the Commission currently does not appear to favor the "no economic sense" test in the case of a refusal to supply, but instead imposes a particularly high burden on dominant companies that even goes beyond showing that the status quo would be maintained following the termination of an existing supply arrangement.

77 Damien Geradin, *supra* note 72.

78 In the *Aspen Skiing* case, in which the courts prohibited a refusal to deal, there had been previous dealings between the parties. The refusal to supply ski passes even at retail prices was clearly intended to harm the competitor. *See* Aspen Skiing Co. v. Aspen Highlands Skiing Corp., 472 U.S. 585 (1985).

79 Discussion Paper, *supra* note 9, at para. 224.

As noted above, Commissioner Kroes has also indicated that an efficiency defense will be considered in Article 82 cases in appropriate circumstances. The Discussion Paper[80] deals with efficiencies as part of the objective justification criterion and, therefore, as a constituent element of the finding of an abuse. In particular, according to the Discussion Paper, it may be possible to justify a refusal to license an IP right if exclusive use of the right is required in order to ensure that the company can recoup the investment it has made in creating that IP right. Consequently, within the framework of the efficiency defense, it is necessary to show that the refusal to deal with others was indispensable to the initial investment. Conversely, the Discussion Paper indicates that a refusal is more likely to be abusive if the investment that led to the indispensable input would have been made even if the investor had known that it would have a duty to supply.

It is hard to see how efficiency arguments could play a role—outside the considerations about the necessity to recoup the investment required to obtain the IP right. However, if there are exceptional cases where efficiencies would be important in a context other than to recoup the original investment, the IP right holder would have to satisfy all of the conditions that were identified by Commissioner Kroes in her speech at Fordham and that are identical to the conditions which must be satisfied under Article 81(3). In summary, the introduction of an efficiency defense in Article 82 enforcement is, in practice, unlikely to change the incidence of cases where the Commission orders the compulsory license of an IP right.

In the *Microsoft* decision, the necessity to withhold interoperability information from competitors for efficiency reasons played an important role. Microsoft claimed that the disclosure of the information in question would seriously damage incentives to innovate. In dismissing this as an objective justification, the Commission took into account the fact that the disclosure of the information would not allow others to free ride on Microsoft's investment by copying Microsoft, and only would give them sufficient information so that they could design products that could interoperate with Microsoft's products. The Commission also weighed the compulsory license's negative impact on Microsoft's incentives to innovate against its positive impact on the level of innovation across the whole industry in determining that the refusal was not objectively justified. In line with the *Microsoft* decision, the Discussion Paper suggests that for interoperability information "it may not be appropriate to apply to . . . refusals to supply [such] information the same high standards for intervention" as those applied to the obligation to license IP rights.[81]

80 *Id.* at paras. 224 and 235.

81 *Id.* at para. 242.

E. EXISTENCE OF A SECONDARY MARKET

With regard to either IP rights or physical property, there is no obligation on a dominant company to supply a product that a competitor simply wishes to resell. The obligation to deal can only arise in circumstances where access to the property or supply of the license or raw material is necessary to compete in a separate market and where it leads to a negative effect on competition in the downstream market.

The Court in *IMS Health* clarified that, in order for there to be an obligation to supply, two distinct markets must be involved—the one in which a dominant position is held (the upstream market) and the secondary market in which the company requesting access wishes to compete in (the downstream market). The Court confirmed that a dominant company is not obliged to supply a product to a competitor for that competitor to simply resell. However, the Court may have considerably reduced the hurdle of finding the existence of two distinct markets by stating that this could be a "potential market or even a hypothetical market." In IP terms, this statement may be extremely broad as an IP right could potentially always constitute a "hypothetical market" and is nearly always used as an input in the creation of an output.[82] Similarly, the interoperability information in question in the *Microsoft* decision may meet this very low test of a "hypothetical market" as there is clearly demand for the information.[83]

The Discussion Paper recognizes the specificities of IP rights by acknowledging that there is no general obligation for the IP right holder to license the IP right. Even where the holder acquires a dominant position, there is no obligation because it is the very aim of the IP right to exclude others from using the IP right to produce and distribute products without the consent of the holder of the rights.[84] It also explains the distinction between the upstream market and the downstream market.[85] However, it does not provide further guidance on the interpretation of the term "potential market or even hypothetical market" used by the Court in *IMS Health*.

F. NATURE OF THE IP RIGHT

Against the background of past cases, the question has risen whether the standards to be applied by a competition authority in relation to compulsory licensing

82 Damien Geradin, *supra* note 72.

83 The Commission took care to point out that the information would not allow the recreation of Microsoft products but would merely allow the development of compatible products.

84 Discussion Paper, *supra* note 9, at para. 238.

85 *Id.* at para. 208.

will differ according to the importance of the relevant right. It cannot be excluded that the Court in *Magill* and *IMS Health* was motivated to use competition law to remedy a situation that it felt to be unsatisfactory under IP law. Accordingly, it is not clear if the questionable nature of the copyright in *Magill* and *IMS Health* and the degree to which such information was worthy of protection, influenced either the Commission or the EC Courts in finding that the holders of those rights were obliged to license to third parties in certain circumstances.[86]

While one may question if it is not the task of IP law, rather than competition law, to ensure that only information worthy of protection is the subject of an IP right, the nature of the IP right may influence a competition law regulator's willingness to interfere with the IP right.[87] In a preliminary ruling on the *Microsoft* decision, the President of the CFI noted that the extent to which the information in question is known or secret is a relevant factor to be taken into account as well as the possible relevance of the value of the information concerned.[88]

In the *Microsoft* decision, the extent to which the information concerned IP rights was left open by the Commission. However, the Commission made it clear in its decision that it was not requiring disclosure of Microsoft's source codes so that third parties could copy Windows, but of the interface specifications so that compatible products could be developed. The Commission explained that interface specifications describe "what" an implementation must achieve—not "how" it is achieved.

The use of competition law to remedy a situation that perhaps could be better dealt with by IP law is also illustrated by the recent *AstraZeneca* decision, which is currently under appeal to the EC Courts.[89] In that case, it was found that a misuse of the patent system was an abuse of *AstraZeneca's* dominant position under

86 *Magill* concerned the copyright in weekly television program listings in Ireland and the United Kingdom, which outside those jurisdictions were unlikely to have been protected by copyright laws on the grounds of lack of originality. Similarly, in *IMS Health*, the degree of creativity involved in the information protected by copyright is questionable and was granted as a result of EC-wide legislation on the protection of databases which provides a very low standard for the application of copyright protection. The *IMS Health* case concerned a refusal to license the right to use the "brick structure" developed by *IMS Health* for processing data received from pharmaceutical wholesalers, and which, according to the complainant, was indispensable for the provision of data on the sale of pharmaceutical products to pharmaceutical companies.

87 *See* Presentation by Carl Baudenbacher, *The protection of intellectual property in a system of free competition,* Studienvereinigung Kartellrecht, Brussels, Apr. 7, 2005 (on file with the author).

88 Order of the President of the Court of First Instance, Case T-201/04R, Microsoft v. Commission, Dec. 22, 2004.

89 Case T-321/05, AstraZeneca v. Commission [hereinafter *AstraZeneca*].

Article 82. One wonders, however, if it would be more appropriate to make provision within IP law for the punishment of companies that misuse the system or provide misleading information so that action can be taken against all companies, regardless of market power, who engage in such practices. Similarly, where an abuse is the result of a loophole in the patent law, then perhaps that loophole should simply be closed (as has now been done in relation to the loophole that had been used by AstraZeneca) rather than action be taken by competition regulators.

One wonders, however, if it would be more appropriate to make provision within IP law for the punishment of companies that misuse the system or provide misleading information so that action can be taken against all companies, regardless of market power, who engage in such practices.

IV. Conclusion on the Impact of the Article 82 Review on IP Rights

Indications so far are that the Article 82 review may not lead to any great change in the circumstances in which compulsory licensing will be considered. However, a deeper understanding of the underlying economic theories of competitive harm may assist competition regulators in distinguishing between pro-competitive licensing restrictions and anticompetitive licensing restrictions.[90]

The reason why the Article 82 review may not impact refusal to license IP rights cases may be because such cases were already limited to "exceptional circumstances" rather than being a per se violation of Article 82 and because the ability to objectively justify a refusal that has anticompetitive effects may have already included an efficiency defense. Similarly, the requirement that the refusal will only be prohibited if it results in the elimination of competition indicates that, in refusal to supply or license cases, EC law takes into account the foreclosure effects of the refusal on the marketplace. This indicates that, in this regard, perhaps EC law already sought (in theory at least) to protect competition as opposed to particular competitors and that the Discussion Paper and the comments made by Commissioner Kroes in relation to the Article 82 review and a more economics effects-based approach, in general, do not indicate a change of policy with regard to first-time refusals to supply or license. However, the weight attached in the Discussion Paper to existing commercial arrangements should be treated with caution and could result in behavior that previously would not have been considered as abusive, falling foul of Article 82. Although not considered in the Discussion Paper, in our view, the "no economic sense test" could be use-

90 In relation to the benefits of a deeper understanding of the underlying economic theories of competitive harm, see Speech by Mark Delrahim, *U.S. and EU Approaches to the Antitrust Analysis of Intellectual Property Licensing: Observations from the Enforcement Perspective*, American Bar Association's Section of Antitrust Law, Washington, DC, Apr. 1, 2004.

ful in determining whether a refusal to continue supplying an existing customer is objectively justified. In addition, the statement in the Discussion Paper that in certain circumstances there may be an obligation to license an IP right where that right is indispensable for the development of a new product that has not yet been identified, should also be treated with concern as this could give rise to compulsory licenses in dubious cases.

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Competition in EU Trading and Post-Trading Service Markets

Bernhard Friess and Sean Greenaway

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Competition in EU Trading and Post-Trading Service Markets

Bernhard Friess and Sean Greenaway

The structure of trading service markets is a fundamental determinant of the cost of capital for business. Competition has an important role to play in delivering efficiencies, particularly in the context of inherited fragmentation that characterizes the European Union, and to this end regulation and competition policy need to go hand in hand. Despite the complexity of the sector, competition authorities need to be alert to the problems that it poses. We argue that competition between trading platforms is welfare-enhancing but often foreclosed, both by private and state measures. In clearing, we take the view that compatibility is needed but unlikely to arise endogenously. In settlement, finally, we tentatively conclude that agency does not influence rents available to central securities depositories (CSDs), but may add value by keeping custody markets contestable.

Bernhard Friess is Head of Unit and Sean Greenaway is Case Handler at DG Competition, Financial Services Unit (D1). This paper only reflects the views of its authors; it does not reflect any official position of the European Commission or DG Competition.

I. Introduction

As in many other areas, the European Union has inherited financial market structures that are characterized both by national infrastructures and a lack of cross-border integration.

Behind these structures lie the financial markets themselves. Financial markets play a key role in the modern economy, both in ensuring efficient capital allocation and in overcoming principal-agent problems for large, diversified corporations by giving transparency to value creation. The free movement of capital not only enhances welfare, but is one of the core freedoms under the EC Treaties. Studies have suggested that integrated financial markets could add the order of one percent to EU GDP, even on a conservative estimate. For this reason, achieving integrated markets is one of the core goals that the European Union has set for itself in the framework of its Lisbon growth agenda.

Firms that organize markets provide services in a market which is obviously distinct from the markets that they organize. Such "trading services markets" typically display network effects, economies of scale and scope, and two- or multi-sidedness. These characteristics may be shared by some of the traditional utilities that have been the focus of past and ongoing liberalization efforts in the European Union, which suggests looking here for inspiration. However, there are also important differences. Unlike in telephony, energy, water, and railways, existing infrastructure operators face low incremental costs in deploying their infrastructure to serve new markets if they can overcome other barriers to entry. The services provided across this infrastructure are also inseparable from the infrastructure itself. While various services can be provided across telephone and electricity networks and it is possible merely to operate the infrastructure, an exchange offers a complete value proposition in terms of its market model, that is hardwired into the infrastructure design. Finally, most exchanges and post-trading infrastructures historically have been mutual organizations, while there has been a more recent trend towards demutualization, particularly at the exchange level—state involvement is significant in terms of regulation, but rare in terms of ownership. This is, of course, similar to the situation of utilities in the United States.

Because the traditional utilities offer commodity products and require a local presence to provide local delivery, cross-border demand is purely wholesale in nature. Thus, no consumer demands foreign energy or water as such, and consumers demand foreign telephony and railways only in order to reach people or places located in (or beyond) the corresponding foreign territory. In order to solve this problem, traditional utilities negotiate terms to access each others' network. By contrast, consumers of financial services regularly seek to trade instruments that can only be traded on foreign infrastructures, even if this demand is significantly attenuated by the high costs of cross-border clearing and settlement. Because there is usually no home network either with which to negotiate wholesale access fees or that is demanding access, consumers use a variety of costly workarounds: essentially they would need to incur sunk costs in every market while the volume of their activity in that market may not justify this. These sunk costs are often imposed on prudential grounds by foreign infrastructures, or form part of the pricing model those infrastructures have selected. Because of this, the use of intermediaries to access foreign markets, that can spread the fixed costs over a wider base, is common (it should, however, be noted that a central securities depository (CSD) can sometimes act as an intermediary in this way).

Before we analyze the problems the industry poses, we describe, in broad lines, how it operates.

II. The Role and Functioning of Trading and Post-Trading Service Markets

A. TRADING AND POST-TRADING SERVICES GENERALLY

The problem solved by the trading and post-trading services industry is simple and archetypal: how to allow potential sellers and buyers of a given instrument to trade that instrument between each other at least cost. Financial instruments confer title, or the right to obtain or abrogate title at a given price, to financial assets such as company equity, company and government debt, and currencies, or commodities such as oil, aluminum, or wheat. Because of their commodity nature, markets in such instruments can be extremely efficient. However, there are also many systems that resemble them and that might be similarly analyzed:

- those that exist to enable trading of less commoditized instruments in respect of which information asymmetries and transport costs may be important (for example eBay);
- those that trade rights, such as carbon dioxide emission rights;
- those that allow hedging against non-financial future events, such as the weather or political outcomes;
- wholesale trading systems in fields like insurance; and
- personal networking schemes.

The common feature of all these systems is that they reduce the search and contracting costs faced by persons wishing to enter into a certain type of transaction. In this way, they are no different, in principle, from the organizer of a mediaeval marketplace and many other physical markets today. In such a marketplace:

- Traders come together to trade among themselves or with the public, and the public comes in order to trade, thereby realizing economies of scope (reducing search costs on both the demand and the supply side);

- The terms of trade may be wholly or partly regulated by the marketplace operator. It may offer additional guarantees to those offered by the trader, thereby underwriting the trade risk, it may offer a mediation service, which may even be binding, or it may undertake to expel unreliable traders. In these ways, it reduces contracting cost on the demand side, namely the risk of adverse selection (obtaining inferior quality due to information asymmetries) and of failure to conclude the trade.

- Demand-side contracting cost may also be reduced endogenously in such a marketplace due to reputational considerations (interactions are repeated and reputations built rapidly and efficiently).

- The marketplace operator also might underwrite the credit risk, especially in an inter-dealer market, if trades are not settled immediately, thereby reducing contracting cost also on the supply side. By doing so, it also adds liquidity because dealers can trade during the day without worrying about their net cash position. More trades occur, therefore more value is created, and the operator can capture more of this value for itself. However, it also generates a risk that unscrupulous, reckless, inexpert, or even simply unlucky dealers are unable to settle trades at the end of the day. Traders may gain collectively from this facility, but individually they will want to offload this risk that the marketplace operator must then underwrite.

Recall that risks matter because risk has a cost, and therefore, leads to a spread between prices offered and bid. The existence of this spread discourages some potential traders from trading, namely all those who value the instrument within the spread. A wider spread reduces the overall volume of transactions offered and concluded on the market, meaning that the fixed costs of running the market increase in per-unit terms, leading to a vicious circle as even higher costs reduce trading further. Similarly, there is a virtuous circle in the other direction, meaning that trading platforms may have an incentive to subsidize the supply of liquidity (as we shall see shortly).

The components of the spread are as follows:

- Transaction costs, both traders' costs of trading and the cost of the market infrastructure itself, known as market friction;

- The premium required to assume adverse selection risk (i.e., having traded when ex-post one would wish to have avoided doing so); and

- The premium required to assume inventory risk (this is the flipside of adverse selection risk, namely the risk of the market moving against a position held by dealers before they can trade out of that position).

B. SPECIFICITIES OF TRADING IN CASH SECURITIES

In what follows, we consider only so-called cash markets, which have been the focus of DG Competition's work so far. These are the markets for immediate transactions in equities and bonds (despite the name, they have nothing to do with currency). While infinitely many option positions can be constructed by any third party, cash trading relies on the presence in the market of underlying instruments issued by firms or governments. These instruments are present in the market only if they have been sold in a given quantity to initial investors, in what is termed the "primary market." Subsequent trading in these instruments between investors has no direct financial consequences for the issuers (at least until individual equity holdings reach levels that have consequences for governance). However, the cost of trading in the secondary market is anticipated by purchasers in the primary market, giving rise to what is termed an "illiquidity discount." As a result, firms raise less from their bond and equity issues than they otherwise would, meaning that their cost of capital is increased. When the cost of capital rises, economic activity contracts as marginally profitable projects are abandoned. Within the European Union, at least, this is no insignificant phenomenon. Estimates are that the illiquidity discount raises the cost of capital for listed firms by about 2.5 percent in relative terms, or the cost of equity by 0.5 percent in absolute terms.[1]

Since dealers in open markets make irrevocable offers to trade on trading platforms, and not only in relation to clients, their prices on-exchange already reflect their costs and inventory risks, but the prices that they quote to final investors may include a further markup. Final investors themselves, of course, internalize transaction and inventory costs in their turn. Dissuaded from implementing what would be profitable investment strategies in a frictionless world, this is the ultimate source of illiquidity.

As mentioned, illiquidity, represented by the bid/ask spread, has various components of which transaction costs are only one. This means that market quality and so-called microstructure (trading rules such as tick size and order preferencing) cannot be neglected. However, the transaction costs in the bid/ask spread, especially as faced by final investors, constitute a very significant part of the total. These costs arise partly in the infrastructures themselves (in the form of fees charged to brokers), but mainly in the brokerage layer itself.[2] Moreover, since the needs of final investors are the reason why markets exist, it would be wrong to view illiquidity costs as only arising at the wholesale level. The reasons

1 London Economics, *Quantification of the macro-economic impact of integration of EU financial markets*, Report to the European Commission (Nov. 2002), *available at* http://europa.eu.int/comm/internal_market/securities/docs/studies/report-londonecon_en.zip.

2 Eurogroup Consulting, *Analyse comparative du coût des operations sur titres en Europe et aux USA*, Report commissioned by the international group AFTI (Nov. 5, 2002), *available at* http://www.telekurs.fr/lesplus/downloads/Etudecomparative/01-Rapport_COUTS_VF.pdf.

for high brokerage costs are not fully clear, but they certainly include the costs of intermediation in foreign markets and of maintaining duplicate infrastructure and links and, thus, partly reflect and magnify the inefficiencies existing at the infrastructure level.

So what does this industry look like? It is important to bear in mind the regulatory and self-regulatory environment of exchanges, which exists in order to limit the risk of default and/or insider trading on financial markets, but which can have competition implications.

Exchange membership is limited by the exchange in order to manage this risk. For this reason, individuals and most investment firms have no direct access to the exchange (although systems provided by their brokers may look like direct access, trading on such systems is obviously independently priced). Similar restrictions arise in clearing and settlement.

However, investors do not only trade through intermediaries because they are excluded from exchange membership. Exchanges do not own the shares that they trade, and by themselves are as useless as a marketplace without any stallholders. Dealers arise in these markets because investors (and hence also issuers) value liquidity, that is to say, the ability to execute trades immediately. However, orders to buy and sell do not arrive simultaneously and would not arrive at all if the investor could not be confident of the price she would pay, at least within a certain range. Because of asynchronous supply and demand, the ability to execute immediately is only available if some market participants act as buffers and hold securities as inventory. Such dealers, however, usually have no fundamental knowledge of the value of securities and no wish to hold them per se. On the contrary, they are exposed to an inventory risk because prices may move against their positions. For this reason, they must discover prices that balance supply and demand in the short term. The exchange facilitates this and allows dealers to manage their inventory positions. As an inter-dealer market, it adds liquidity by reducing inventory risk. When a dealer wishes to dispose of an inventory position on the exchange, or acquire such a position in order to meet its obligations, it may demand liquidity.

IT IS IMPORTANT TO BEAR IN MIND THE REGULATORY AND SELF-REGULATORY ENVIRONMENT OF EXCHANGES, WHICH EXISTS IN ORDER TO LIMIT THE RISK OF DEFAULT AND/OR INSIDER TRADING ON FINANCIAL MARKETS, BUT WHICH CAN HAVE COMPETITION IMPLICATIONS.

Dealers, then, interact with investors and with each other in what we term the "market for liquidity." In this market, demand comes from both dealers and investors, whereas only dealers can cost-effectively supply. This can be likened to the market for antiques, in which private persons may wish to acquire or dispose of items, but they cannot offer to do so directly because of foregone economies of scope (high search costs) and the adverse selection risks they would incur or impose on counterparties. Dealers also wish to acquire items for sale against known customer demand and dispose of items in their inventory for which local demand is weak. Therefore, they supply private investors with liquidity while competing against them by demanding liquidity. This competition, of course, is potentially one-sided because private investors often do not have access to inter-dealer markets and certainly do not have such access in the case of securities. Exchange rules endeavor to manage this conflict of interest.

Exchange members are not only dealers. Some members never take a position in anything and simply pass orders on to dealers or to the exchange. These are known as brokers. Finally, some players with investment, speculative, hedging, or arbitrage motives do have direct access to the exchange—typically these are banks who are also broker-dealers.

Since membership is restricted, we are led to the conclusion that trading services markets are effectively segmented into fully distinct markets from the demand standpoint. What we can call the "central" trading services markets provide these services to the members of the trading platform or of post-trading infrastructures. There are then "peripheral" markets that provide trading services to institutional and private investors, as well as to those potential consumers of central trading services who are not served in that market and, therefore, seek intermediated access. The peripheral market is dependent for its existence on the ability of the central market efficiently to form prices and pool risk. The exchange, or any other trading system in the central markets, brings together suppliers and demanders of liquidity. By contrast, automated trading systems in the peripheral markets, where they exist, are operated by suppliers of liquidity and serve demanders alone. Although online order entry is increasing, the peripheral markets still make extensive use of manual systems to place orders.

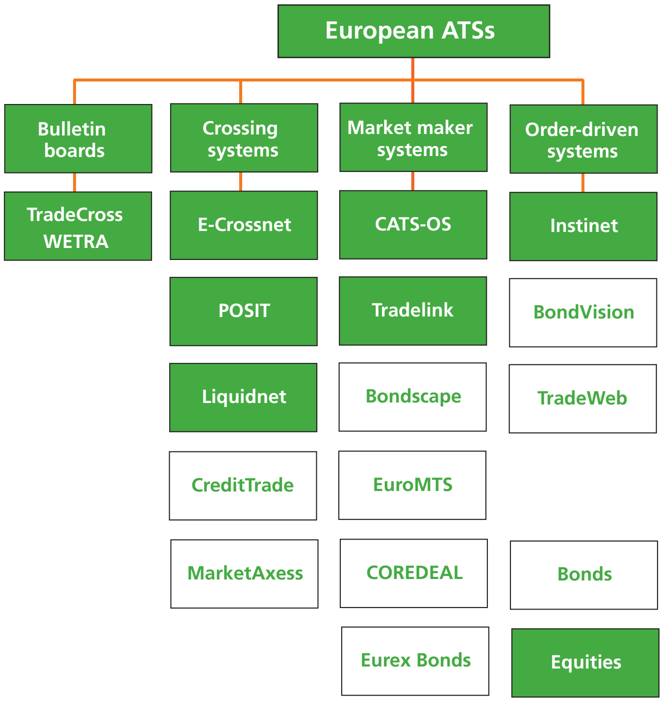
III. The Organization of EU Trading and Post-Trading Services

A. EXCHANGES AND ALTERNATIVE TRADING PLATFORMS

Trading platforms can be divided into regulated markets and multilateral trading facilities (MTFs)—also referred to as Alternative Trading Systems (ATSs). The difference between regulated markets and MTFs is one of regulation. Although there are many important differences between the market models implemented by exchanges, these days they almost always operate using an electronic order book. The principal characteristic of this system is that counterparties cannot be selected by suppliers of liquidity, whose offers are free to be hit by any member of the exchange. There is more variety in the market model of MTFs, ranging from mere bulletin boards that do not arrange trades at all, to crossing systems that trade large blocks at prices derived from the exchange, to fully-fledged alterna-

Figure 1

European MTFs



Source: Marion Muhlberger, Alternative Trading Systems: A Catalyst of Change in Securities Trading (Deutsche Bank Research, Economics: Digital economy and structural change No. 47, Jan. 11, 2005).

tive order books with price discovery. In addition to a possible price advantage over the exchange, such systems may have other advantages, such as after-hours trading, ease of trading larger blocks of shares, and alternative or more flexible settlement arrangements. In the equity arena, MTFs are believed to have limited market share (although data on this is lacking), but they are much more important for bonds. Many exchanges also operate what are technically MTFs with different or hybrid market models for less liquid shares.[3]

The most important EU exchanges are:

- The London Stock Exchange plc (LSE);
- Deutsche Börse AG (DBAG), which operates the Frankfurt exchange as well as the Eurex options exchange; and,
- Euronext nv, which operates the Paris, Amsterdam, Brussels, and Lisbon equity markets as well as the London international financial futures exchange (Liffe).

3 The list of regulated markets pursuant to the EU's Investment Services Directive is kept by the Commission and available on its website.

Each EU member state has a cash equity exchange and several have more than one. In the next tier by size, we have the Spanish and Italian markets and the Scandinavian markets operated by OMX.

Technically, the most important MTF is probably EuroMTS, operator of EU bond markets, which is only a regulated market for its Italian and Portuguese government bond segments. Others are shown in Figure 1.

B. MANAGING DEFAULT RISK AND THE ROLE OF CENTRAL COUNTERPARTIES

Once a trade occurs—that is to say, once two counterparties are matched on the trading platform—it is necessary to ensure that the actual securities involved in the trade are exchanged and payment takes place. For various reasons, no trading system offers real-time gross payment (i.e., immediate payment and immediate delivery of the corresponding security each time a trade occurs). Typically, the obligation actually to exchange the securities and make payment only arises three days after the trade has been entered into. This arrangement serves, in particular, to add liquidity by relaxing cash flow constraints, but it creates a risk of defaulting on trades. It may be that during the settlement period one of the original counterparties has become insolvent and so is unable to honor her commitment, or it may be that at the moment of delivery the selling party does not have the security in question, or the buying party does not have sufficient cash at hand, or one of the parties has an incentive to default because of price movements in the meanwhile. If this were to happen, it would severely damage confidence in the market. Therefore, exchanges are under a statutory requirement to ensure that trades can be expected to clear and settle, and other trading platforms have a similar incentive.

The major exchanges have responded to this challenge by introducing a central counterparty (CCP) into their market model. The CCP reduces the risk of default by interposing itself on both sides of each trade, so that it guarantees settlement. In other words, it becomes the seller to every buyer and the buyer to every seller. A CCP has other advantages too. By reducing the number of bilateral exposures by a log factor of two, it is able to net positions much as a payments clearinghouse does, meaning that many fewer settlement operations need to take place. By interposing itself in every trade, it also allows market participants to retain their anonymity relative to each other, which is a positive feature of market design because it aids liquidity provision and reduces volatility (this is because it disables inferences on adverse selection from the identity of the counterparty). A CCP does need to call for collateral to cover the positions to which it is exposed, but in this respect it also realizes economies by allowing offsetting positions to be netted. This is particularly significant when it can operate across cash and derivative markets or markets for other asset classes. The CCP does have consequences for competition in the trading services market, which we will discuss shortly. Most ambiguous is whether the CCP actually provides a service to the exchange or to its own clearing members. These are a subset of exchange members, since some members may choose to clear (and settle) through an intermediary.

A CCP is not necessarily limited to providing services only to an exchange—it can also be an obligatory or optional part of trading on other platforms and be used for bilateral trades. When it does provide services in the same security for trades realized on different platforms, the question of full fungibility arises (i.e. whether the positions of a single member on both platforms can be offset against each other to produce a single collateral position and a single position for settlement). If they cannot, the attractiveness of the second platform may decline. However, in practice, CCPs may be constrained in providing fungibility by conditions in their contracts with exchanges or by the exercise of direct control.

In the European Union, CCPs are vertically integrated with the exchange in Italy and effectively in Germany (Eurex Clearing), whereas there is no CCP in Scandinavia or in Spain. The only independent player in the European Union, LCH.Clearnet, has minority exchange ownership and is otherwise owned by banks. In fact, LCH.Clearnet consists of two CCPs, one serving the LSE and the other the continental markets. LCH.Clearnet also clears for MTS. While a for-profit corporation, there are limits on the profits it can retain.[4] The Depository Trust & Clearing Corporation (DTCC), the U.S. clearer, is also sometimes cited as a possible competitor for EU CCPs. DTCC is vertically integrated with settlement.

C. CENTRAL SECURITIES DEPOSITARIES

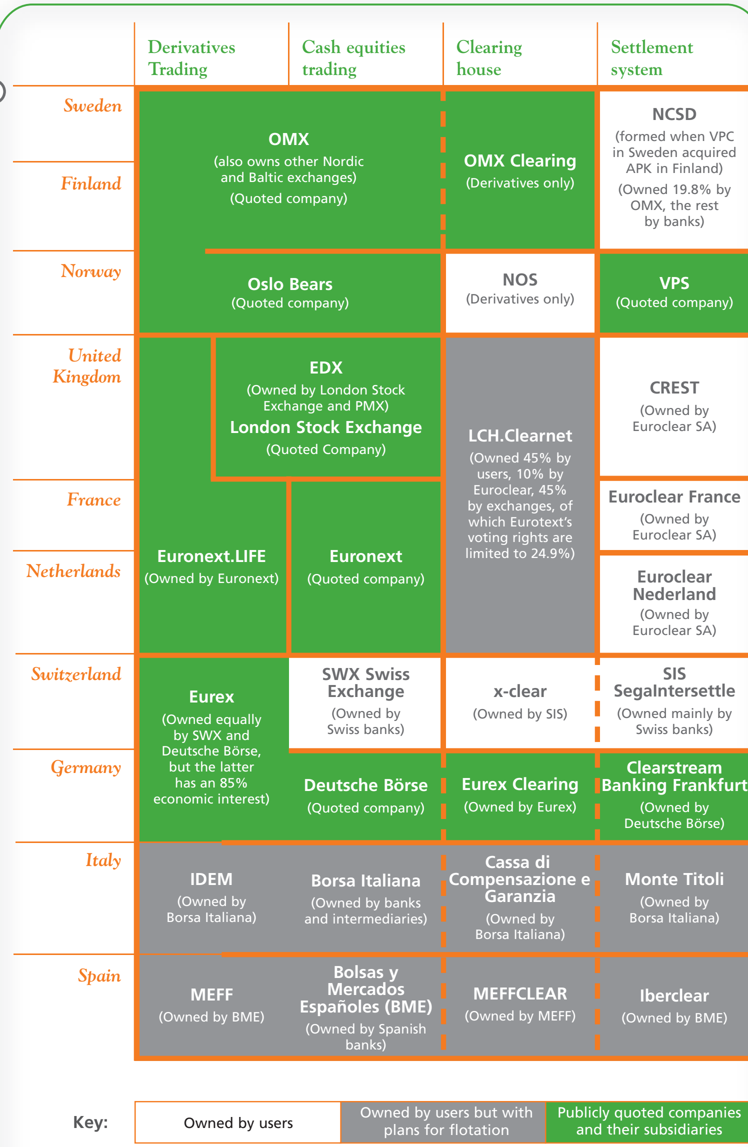
Regardless of the trading platform, eventually all positions accumulated that cannot be netted against each other have to be settled. In equities, this implies an irreplaceable role for so-called Central Securities Depositories (CSDs) owing to the need to keep track of ownership. In Eurobonds, the CSD functions are generally provided by the two so-called International Central Securities Depositories (ICSDs), Euroclear and Clearstream, which settle in commercial bank money, provide lending facilities to guarantee settlement, and also offer custody services. The certificates of deposit behind bonds are normally lodged with a national CSD when issued for trading on exchange, or with so-called "common depositories" (banks) in other cases. These entities also manage any changes in the net positions of the two ICSDs. The CSD for government bonds is occasionally the central bank.

Settlement is inseparable from custody, but custody has to take place even if there is no trading. This is because the owner of a bond may need to be traced in

4 Once LCH.Clearnet's EBIT exceeds 150 million in any given year, 70 percent of this excess is to be for the benefit of users, in a manner to be determined by the LCH.Clearnet Board. *See* Announcement, Euronext, Clearnet and LCH to merge to form the LCH.Clearnet Group (Jun. 25, 2003), *at* http://www.euronext.com/vgn/images/portal/cit_88313/16/33/377913248LCHCLEARNET_en.pdf (last visited Feb. 9, 2006).

Figure 2

Overview of principal European financial market infrastructures



Source: Adapted from Bourse Consult, The Future of Clearing and Settlement in Europe (Corporation of London, City Research Series No. 7, Dec. 2005).

order to credit the holder with a coupon payment or the repayment of the principal on maturity. The same applies to dividend payments to equity holders, with added complications such as stock splits, takeovers, and voting rights. Even if an instrument is not traded, someone owns it and so this function needs to be assured.

An individual exchange trade is not necessarily settled in the CSD or an ICSD because of settlement netting. Settlement netting is implied by the presence of a CCP, but it can also be influenced by the use of an agent for settlement. As mentioned, not all trading members of the exchange are necessarily clearing members of the CCP, and the same applies to settlement. If an agent is used, its own trades may be netted in the CCP with those of the parties for whom it is acting as an agent and only a net position will be settled. From the standpoint of the trader using a settlement agent, her net trades effectively settle in the books of that agent rather than in the CSD, but may require the agent itself to settle with the CSD.[5] Settlement is a necessary corollary of any trade and not only of trades on-exchange.

An overview of all the European Union structures we have discussed is provided in Figure 2.

Euroclear is owned by banks and in turn owns the equity CSDs of France, the Netherlands, the United Kingdom, and Belgium. Clearstream is owned by DBAG and in turn owns the German equity CSD, Clearstream Banking Frankfurt (CBF). The settlement infrastructure is also vertically integrated with the exchange in Spain and Italy, while the exchange only has a minority stake in the Scandinavian equity CSD holding VPC (operated commercially under the label NCSD).

IV. The Economics of Trading and Post-Trading Services

A. TWO-SIDEDNESS OF TRADING SERVICES MARKETS

From the brief discussion of industry organization above, we see that the issue of complementarity in the value chain trading - clearing - settlement (- custody) is anything but simple. A further complication is the existence and implications of two-sidedness.

Rochet and Tirole (2005) define a market as two-sided if two groups of customers can be identified such that the volume of transactions realized is sensitive to the division of the price between them.[6]

Several authors have suggested that an exchange is such a two-sided market as between buyers and sellers of securities. In our view, however, this is incorrect. Buyers and sellers do meet on exchanges, but they have the same demand function. No exchange, to our knowledge, offers different fees to buyers as opposed to sellers—nor does any broker. Of course, there is a spread in prices, but there is also a

5 In the majority of cases, CSDs separate own account from client account for settlement agents.

6 J.-C., Rochet & J. Tirole, Two-Sided Markets: an Overview, (U. Toulouse, IDEI Working Paper, 2005), *available at* http://idei.fr/doc/conf/tsm/rochet_tirole.pdf.

pass-through mechanism If the price charged on one side of this putative two-sided market were to change, it would be immediately reflected in the price of the security. This means that, if buyers and sellers are the two sides, then the Rochet-Tirole definition of a two-sided market is not met, because price structure is irrelevant.

Trading platforms, however, do operate in a two-sided market, but the relevant distinction is between suppliers and demanders of liquidity. It is important to be clear that supplying liquidity is not the same thing as offering securities for sale. Because securities, unlike antiques, are commodities, it is possible also to offer a firm price for purchase and not only for sale. Therefore, suppliers of liquidity offer "two-way" prices. Similarly, those who demand liquidity may want to buy or to sell. Here also, whether they want to buy or to sell is not important.

The relevance of this distinction arises from the fact that an unconditional offer either to buy or sell on exchange creates a free option in the market, whereas accepting the offer cancels that option. This is a real difference. The option to buy benefits everyone, but it is not possible for the provider to charge everyone for it. Entering an order or displaying a quote also has an administrative cost, which (unless the supply of liquidity is directly remunerated) can only be recuperated on profits made on orders that are subsequently "hit" (i.e., that find a counterparty). For this reason, there is a consumption externality and the price structure is, in fact, relevant to the equilibrium level of transactions. Suppliers of liquidity must receive a price that subsidizes the option value or liquidity will be undersupplied. Through the platform, demanders of liquidity jointly subsidize its supply.

Note that the situation we are describing is, in a sense, perfectly mundane. The standard retail practice of attaching a price tag to products offered for sale, and indeed offering them for sale at all, represents a concession of option value to consumers (although it, of course, may have countervailing efficiencies). In a store, the retailer is vertically integrated with the "marketplace provider", itself, so there is no need to address the incentive to supply goods. In other situations, such as shopping malls or trade fairs, it is also likely, as an empirical matter, that liquidity supply does not need to be subsidized either because there are offsetting externalities or because competition for demand competes away any attempted subsidy to the supply side. In the trading world, however, market power combined with the market structure does make this relevant.

ANY ANTITRUST ANALYSIS OF PRICING MUST AVOID FALLING INTO THE TRAP OF CONSIDERING EITHER SIDE OF THE MARKET IN ISOLATION FROM THE OTHER.

This has two consequences. One, obviously, is that any antitrust analysis of pricing must avoid falling into the trap of considering either side of the market in isolation from the other. This, as has been frequently pointed out, would be an error of analysis. The second consequence is more subtle, since it has to do with determining the optimal pricing strategy under various vertical structures—the components of which display demand complementarities.

B. COMPLEMENTARITY IN THE VERTICAL CHAIN

In order to assess the efficiencies associated with vertical relationships, it is necessary to discuss the complementarities that exist between trading and post-trading services. However, this subject is far from trivial.

1. Complementarity between Trading and Central Counterparty Clearing

Where a CCP exists, its services are a perfect and one-for-one complement to exchange trading from the standpoint of demanders of liquidity. However, liquidity suppliers will consume the CCP's services in a lower proportion, namely only to the extent that the trades they offer are matched. Thus, one side of the trading market consumes these services in varying, rather than fixed, proportions. This means that the CCP's pricing has a different effect on demand on the two sides of the market, and implicitly it contributes to the balance between the two sides that needs to be considered by the exchange in setting an optimal price.

In addition to the problem of varying proportions, when a CCP receives a matched trade from an exchange, it has, as a practical matter, no way of telling which party to the trade first offered to trade, and as a result, it cannot price-discriminate on this basis. Therefore, it must charge the same price to the two exchange customer groups, regardless of the fact that it effectively is charging liquidity suppliers less on a per-transaction basis. Of course, the exchange feed to the CCP could include this information if it wanted to, but to the best of our knowledge no CCP differentiates its prices in this way. This means that the two-sidedness extant at the trading level does not extend to clearing, nor, by extension, to settlement, and these layers are unable to take it into account.

2. Complementarity of Settlement Services

Settlement services, as we have shown earlier in this paper, are complementary to trading and clearing, but in a variable proportion to both. Settlement providers are unable to internalize the distinction between supply and demand of liquidity, but in addition, when there is settlement netting in the CCP, the settlement price affects traders differently depending on the intensity of their trading during any given clearing period. Once again, the CSD normally has no means to internalize this difference unless it can first-degree price discriminate. However, we note the situation with Crest in the United Kingdom (part of the Euroclear group) which, because it has a role in clearing, is able to, and in fact does, price settlement on a gross (i.e., pre-netting) basis.

C. IMPLICATIONS OF GOVERNANCE

Finally, governance arrangements may influence the extent to which actors at each level may effectively seek to profit-maximize.

In a static monopoly setting, it is clear that perfect user governance can eliminate rents, essentially because any rents achieved are rebated as dividends to users in proportion to their use, and thus, have no net economic effect. The effective price is the billed price less the dividends. However, it is obvious that the lack of a genuine profit motive also has costs associated with it. These include classic elements such as a lack of incentive to innovate, persistent X-inefficiency, and absent incentives to engage in industry consolidation, as well as more subtle elements such as the claim that a user-owned CCP does not have an incentive to encourage the socially optimal level of trading because it imports risk-aversion from its owners.[7] A further element to consider is the ability of a subset of users to exploit ownership of essential facilities to raise rivals' costs, thereby, at a minimum, impeding entry, and potentially underpinning a cartel. There is also increasing awareness of the difficulties that user-governed entities have to overcome principal-agent problems, not only because of the difficulty of aligning agent incentives, but also because principals themselves face a coordination problem and will be tempted to free-ride given that, individually, none of them internalizes the full benefit of their governance efforts.

These elements mean that we need to be cautious when it comes to prescribing user governance as a remedy to market failures. It should be recalled that exchanges come from a background of mutual ownership in which markets were perfectly segmented. Almost all positive developments in the industry since then can be tied to the efforts of private firms to seek out and capture new sources of value. Turning the clock back is not a self-evident strategy.

D. EFFICIENCY OF VERTICAL ARRANGEMENTS

All of these characteristics of the market need to be kept in mind in considering whether there are intrinsic efficiency problems in the trading chain and, if there are, whether vertical integration or other vertical contracting relationships solve, or potentially exacerbate, the problems.

In addition to the classical issue of double (or triple) marginalization, we also need to consider the impact of complementarity on the two-sidedness faced by the trading platform, recalling that this two-sidedness is fundamental since it impacts the supply of liquidity.[8]

7 T. Köppl & C. Monnet, Central Counterparties (May 2005) (preliminary draft), *available at* http://www.cass.city.ac.uk/conferences/clearingandsettlement/koeppl_monnet.pdf.

8 This paper does not consider whether vertical integration realizes economies of scope or scale or efficiently solves technical contracting problems. However, it seems probable that any gains of this kind would be comparatively minor in comparison to gains from more efficient pricing.

1. Monopoly Rents

The first point to emphasize is that any market structure is likely to realize monopoly rents, since all three services are complementary to at least some degree, and as long as any single layer is a monopoly that seeks to profit-maximize, there will be monopoly rents in the entire chain. Absent the ability to perfectly price-discriminate, this will further lead to deadweight losses, and deadweight losses in the industry will translate to macroeconomic losses via a multiplier of uncertain, but very significant, size.

It should be pointed out that the estimates of savings from market integration that have been carried out have been based on making cross-border transactions as easy and as cheap as domestic ones are today. This is certainly an important goal to pursue. However, it says nothing about additional gains that might result if the entire market were competitive and realizing only normal profits. We would not expect any market to be realizing only normal profits and to be free of X-inefficiency today.[9]

A second issue which is interesting to consider, is whether, given complementarity in varying proportions, it is worse to have a monopoly in one layer rather than another. It appears that, from this angle, a trading monopoly is likely to be the most welfare-reducing. This is because the players at the other levels, when they set their privately optimal price, impact only a part of the transactions. It should be acknowledged that, in some circumstances, competing two-sided networks might be unable to address two-sidedness as efficiently as a monopolist. However, it seems that this is not true of trading platforms, because they can individually internalize the externality created by the supply of liquidity. Ideally, on both these conjectures, more rigorous analysis should be performed.

2. Double Marginalization

Given pricing interdependencies, it is clear that there is potentially a coordination problem to be solved in the vertical chain. A vertically integrated structure has access to all the information available to optimize pricing, not only through the chain, but also across the two sides of the trading market. This might be difficult to reproduce contractually. The main question is whether, even with this information, the problem is tractable. As an empirical matter, it seems that there are considerable pricing rigidities that need to be explained. Thus, we are not aware of a vertical silo adopting a pricing model that either charges settlement in proportion to gross rather than net trading or differentiates in clearing charges between supply and demand of liquidity. Therefore, it is not obvious that the

9 X-inefficiency refers to the failure to maximize profits due to private benefits which actors within the firm obtain from reduced effort or discretionary expenditure. As a result, observed rates of return on assets may not fully reflect the rents being earned relative to a competitive industry. In essence, some of the market power rents, rather than accruing to shareholders, are consumed as benefits in kind by employees.

overall pricing package has optimized anything more than a non-integrated structure could do.

At the same time, it is conceivable that vertical integration might generate a range of inefficiencies, particularly in the CCP layer. Since CCPs realize economies of scope and scale, their optimum pricing function shifts as they serve more markets. However, a vertically integrated CCP may have less incentive to realize these economies—particularly through consolidation if it would result in weakening of control by the parent company. Similarly, it is a less attractive partner for other trading platforms if it would reduce the cost of the parent exchange challenging in their markets. Finally, as we will discuss in the next section, control of the clearing and settlement layers of a vertical silo offers opportunities for foreclosure of competition in trading.

We also need to bear in mind that user governance may limit the rent-seeking behavior of some clearing and settlement organizations. Where this is so, the problem of double marginalization does not arise.

In light of these considerations, we are skeptical of any net gains from vertical integration imputable to the elimination of double marginalization. This means we need to view efficiencies in the market in a wider framework.

3. Two-Sided Pricing

It is obvious that vertical integration is one way to solve the problem of achieving an efficient two-sided price for trading, in the presence of complementarity with clearing and settlement and market power in those layers. As already indicated, a vertically integrated structure has the advantage, in theory, of having full information allowing it privately to optimize its overall pricing structure.

This problem, however, can also be solved in a non-integrated structure if the trading platform prices last (provided, perhaps, that negative prices on the liquidity-supplying side of the market are possible). In current market structures, an incumbent exchange is unlikely to price last of its own volition, but the same effect could be achieved if the exchange is able to exercise control over clearing and settlement prices and price structures. Moreover, a trading platform that challenges the market is, in fact, likely to price last—it takes the clearing and settlement prices as given. Similarly, if there is competition in trading, then platforms cannot set an overall supra-competitive price, but because there are no cross-platform transactions, they can freely set the pricing mix across the two sides of the market. It follows that, if trading is competitive, the lack of vertical integration does not form an impediment to efficient two-sided pricing.

E. STANDARD-SETTING

A last issue to consider is whether industry structures are such as to make efficient standard-setting arise endogenously. In general, there are two ways standards can arise: through eventual triumph of one standard after a period of competition, and through coordination. The first route is obviously tributary of the degree of competition in the market, and since we are largely interested here in standards as a vehicle for competition rather than outcome of it, we do not need to consider it further. The literature shows that private firms may have an incentive to cooperate on standards in certain circumstances, and may also be able to overcome coordination problems to do so. This result notably arises in network industries when the cooperating parties can internalize a part of the additional network externalities generated by compatibility. It is argued, for instance, that this explains the endogenous origin of computer hardware compatibility (it is also noted that this outcome maximizes social welfare, but may adversely affect consumer welfare when considered alone).[10]

Then the question that arises is whether trading and post-trading markets share similar characteristics such that standards will arise endogenously. As an empirical matter, it does not seem that they do, since little progress has been made in standard setting despite this being identified as a key obstacle to further market integration in the 2001 report by the Giovannini Group.[11] But, on theoretical grounds, we also believe this is not to be expected. The reason is that computer hardware manufacturers compete among each other whether or not there is compatibility. This gives smaller players an incentive to provide compatibility, even unilaterally (assuming this is not impeded by unknown specifications or intellectual property rights). Subsequently, large players have no incentive to reestablish incompatibility unless consumers value it more than the foregone network externalities. By contrast, trading and post-trading infrastructures do not normally compete against each other. Only in the trading-to-user space would there appear to be an incentive for smaller competitors to mimic established interface protocols in order to minimize switching costs for target customers. The same motivation (among others) leads competing trading platforms to seek compatible clearing and settlement arrangements to an incumbent which is, as we will discuss in the next section, more akin to the first route towards establishing standards. But on the whole, we see no economic grounds on which trading and post-trading institutions, at least if they are profit-seeking, will spontaneously pursue shared standards that enable competition, nor do we believe there are any examples of this occurring in practice.

A related question is whether there is a potential market failure in the provision of smart order routing tools. It makes more sense to consider this, however, once we have discussed the conditions determining competition between trading platforms.

10 O. Shy, The Economics of Network Industries ch. 2 (2001).

11 The Giovanni Group, *Cross-Border Clearing and Settlement Arrangements in the European Union* (Nov. 2001), *available at* http://europa.eu.int/comm/economy_finance/publications/giovannini/clearing1101_en.pdf.

V. What Kind of Competition Adds Value and What Prevents It from Occurring?

A. COMPETITION BETWEEN TRADING PLATFORMS

With the adoption of the Markets in Financial Instruments Directive,[12] the European Union committed itself to competition between different forms of trading. Available evidence, although it may be thin, points to benefits from such competition both for fees and for market efficiency:

> "Competition has discovered that diverse traders are willing to support a diversity of trading systems, each of which has evolved to provide low cost service to some constituency . . . . There are strong reasons to believe that the current fragmentation of markets is not particularly costly relative to the service benefits that it provides to diverse clienteles. The widespread availability of market quote and trade data, the ability to route orders to the best prices , and the activities of arbitrageurs all act to integrate fragmented markets."[13]

This is particularly likely to be the case when the fixed costs are sunk because of current market fragmentation, and therefore, do not need to be incurred again in order to offer a competing service. In fact, the marginal cost of launching a competing service may be very low and it may break even on very low market shares.

The question that then arises is whether barriers to entry exist in the form of foreclosure strategies available to the incumbent or from other characteristics of the market.

In this section, we consider two possible barriers to entry: restricted access to the CCP or to the CSD, and the non-availability of market reconstitution technologies. We then close the section with a brief comment on the dominance of incumbents with regard to Article 82 of the EC Treaty.

1. Effective Access to Clearing and Settlement

For competition between trading platforms to develop, a variety of conditions need to be met. These include access to, and fungibility in, clearing and settle-

12 Council Directive (EC) No. 39/2004 on markets in financial instruments (Apr. 21, 2004).

13 L. Harris, The Economics of Best Execution (Mar. 9, 1996, U. S. Cal.) (preliminary draft), *available at* http://www-rcf.usc.edu/~lharris/ACROBAT/BESTEXEC.PDF.

ment. So it is interesting to ask under what conditions CCPs and CSDs will or will not grant this fungibility and whether an incumbent has the means and incentive to foreclose by refusing fungible access.

When the clearinghouse or settlement infrastructure is owned by an incumbent exchange which is under no legal obligation to provide access outside competition law, it does not seem very likely that it will voluntarily do so on non-discriminatory terms. This is because the entrant will compete head-on with the incumbent. Entry is unlikely to grow the market significantly, and competition is likely to be on Bertrand terms, thus destroying industry profits. If the entrant is more efficient, it might even drive the incumbent out entirely.

Admittedly, the incumbent still would control the monopoly post-trading infrastructure, and if the trading segment is competitive, then it could still capture the entire monopoly rent, only incurring the cost of writing off the trading assets. However, this and foregone private benefits are, with high probability, more than sufficient to exclude this outcome from being considered. A corollary of the single monopoly profit is that the vertically integrated structure has no incentive whatsoever to provide access to clearing and settlement, because it cannot boost its profits in any way by doing so.

In conclusion, a vertically integrated structure has both the means and incentive to foreclose. Moreover, even if a new entrant could try to use an alternative clearing provider at a competitive disadvantage, which is difficult enough, it has no chance to avoid the incumbent CSD, while the incumbent has every incentive to concentrate rent capture there. Thus, it is certain that the single monopoly profit will accrue entirely to the incumbent until such time as the entrant has sufficient market share to extract some of it for itself. In a vertical silo, competition in trading is likely to be foreclosed.

It is certain that the single monopoly profit will accrue entirely to the incumbent until such time as the entrant has sufficient market share to extract some of it for itself. In a vertical silo, competition in trading is likely to be foreclosed.

When the CCP is independent and either run for profit or user-owned, it has an incentive to enable entry at the trading level and offer fungibility. This is because, where there is a single player at each level, it is the exchange that typically has the strongest bargaining power in capturing the available monopoly rents. This is seen in both the retrocession fee provided from Clearnet to Euronext and the price auction for clearing services held by LSE. When there are alternative trading platforms, however, the threat of the primary platform replacing the clearing provider is much less existential, since users with a vested interest in not changing clearing provider (e.g. because they use the same provider on other markets or they have incurred high sunk costs in interconnecting systems) can more credibly threaten to switch all or part of their trading to the alternative platform. This alters the bargaining power in the clearinghouse's favor, or in users' favor if users own the clearinghouse. The CCP, therefore, has an incentive to sponsor and facilitate entry.

However, even when the CCP is independent, it may not be free in its actions. This is because it can invite entry only when it enjoys incumbency. However, in order to attain incumbency, it must be awarded it by the exchange. Therefore, the exchange can, and has an incentive to, specify contractual conditions that exclude fungibility or, at the very least, render it more difficult to achieve. The order of moves in the game matters, and the CCP, moving second, is disadvantaged so that an apparent strategy disappears. Of course, it might reappear if the CCP had sufficient power in other markets such that users would switch to it or retain it and start trading on a new trading platform, rather than incur the cost of changing CCP in the first market. But this is not how the market presently works.

On this analysis, it might appear that competition in trading will always be foreclosed if there is a CCP, even if it is independent. This, however, is not true if effective access to the CCP is assured as a consequence of applying competition law. This is either because the CCP, in refusing fungible access, would infringe the dominance provision of Article 82 of the EC Treaty, or because a contractual arrangement with the incumbent exchange preventing the CCP from granting such access might be void under Article 81 of the EC Treaty. If the latter were the case, an independent CCP would then have the means to encourage competition among trading platforms.

As there is little precedent in applying Article 82 in this industry,[14] we limit ourselves to some general considerations. Any obligation to supply pursuant to Article 82 can be established only after close scrutiny of the facts in a given case. This starts from the principle that dominant suppliers, as any undertaking, are usually free to determine whom to supply, and that a refusal to supply may infringe Article 82 only if it has a likely anticompetitive effect on the market that is detrimental to consumer welfare.

Relevant case law normally requires the supplier to have a dominant position in an upstream or related market and to be able to control an input needed to compete in a downstream market.[15] Although a CCP provides services in the first place to its clearing members, it is also a supplier to the exchange since it enables

14 Commission Decision, Case DG COMP/38.096, Clearstream (Clearing and Settlement) (Jun. 2, 2004, not yet reported), *available at* http://europa.eu.int/comm/competition/antitrust/cases/decisions/38096/en.pdf. That decision found that Clearstream, the German CSD owned by Deutsche Börse, had infringed Article 82 by refusing to supply input services to a customer (Euroclear) who competed with a company belonging to the Clearstream group in the downstream market for agent settlement.

15 Case C-7/97, Oscar Bronner GmbH & Co. KG v. Mediaprint Zeitungs- und Zeitschriftenverlag GmbH & Co., 1998 E.C.R. I-7791 and Case C-418/01, IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. KG, 2004 E.C.R. I-5039.

the latter to complete the trades made on its platform. The fact that the London Stock Exchange organized a bidding process for its clearing contract illustrates the existence of a supplier-customer relationship. Even where the CCP is vertically integrated with the exchange and its input is supplied exclusively to its owner, it seems justified to assume the existence of a potential market since demand for that input may arise if a newly competing exchange also requires access.

It would also seem likely that the CCP's input (fungible access) is indispensable for a trading platform to compete effectively with the incumbent exchange downstream. Episodes of competition between exchanges that involved both fungible clearing and remote clearing arrangements appear to confirm that only the former will enable an entrant to compete effectively on price.[16] More specifically, the prospect that an entrant could turn to an alternative supplier or duplicate the CCP is clearly excluded in most cases. Because of economies of scope in clearing, lack of fungibility constitutes a significant impediment to an entrant. While a refusal to grant such access would not completely foreclose a competing trading platform, it is likely to have a significant negative effect on the level of competition in the downstream trading market which, as instances of actual entry show, can reduce trading fees significantly if supported by fungible clearing arrangements.

Similar considerations could be made in assessing an exclusive contractual arrangement between the incumbent exchange and an independent CCP under Article 81. Whether the exchange imposes exclusivity on the CCP or puts in place arrangements that have a similar effect (e.g., a right of first refusal), such arrangements may appreciably restrict competition if they foreclose or significantly impede competition at the trading level.

Any refusal to supply, therefore, would need to be objectively justified by efficiencies, both under Article 82 and under the exemption clause of Article 81. Whether investment incentives could justify the exclusion of an entrant from access to the CCP, at least for a certain period, may have to be considered on a case-by-case basis. However, even within a vertically integrated structure, this is not obvious. The comparison with non-integrated infrastructures suggests that a CCP can achieve significant efficiencies even where it serves different trading platforms.

Of course, it may be argued that competition between trading platforms would be a rather relative gain, since a significant chunk of the monopoly profit previously present in the trading layer would not be competed away, but would merely migrate to the CCP. If the CCP is owned by the community of users in proportion to their use, then this problem would not arise because it will rebate its profit to users. If it is operated for profit, however, then this objection would have more force. If it is owned by a subset of users, then those users can potentially raise rivals' costs by setting high clearing fees that will then be partly rebated to

16 One example is the trading services offered by the London Stock Exchange and Deutsche Börse for Dutch cash equities in competition with the incumbent, Euronext.

the owners in the form of dividends, giving them a cost advantage. In this case, the welfare gain depends on how competitive the market composed of the user-owners is, compared to potential entrants that are excluded. Clearly if this market is perfectly competitive then there are no welfare effects from this strategy (it may even raise welfare by preventing inefficient entry), but if it is oligopolistic then the monopoly rent is indeed recaptured in part.

2. The Provision of Market Reconstitution Technology

Broker-dealers will only look at liquidity furnished on alternative trading platforms if there is a concrete and compelling reason to trade there, or if they can quickly and easily compare quotes across platforms from the standpoint of total trading costs. This implies the existence of a technology layer which reintegrates the underlying fragmentation from a user perspective.

In the United States, best execution requirements on brokers require them to perform this comparison. As a result, this technology exists and is operational there. In other words, it was not necessary to determine if the market would spontaneously deliver it because demand for it was regulated into existence. In EU markets, this kind of technology is rare. Apparently it is a chicken-and-egg conundrum—the lack of such technology dissuades entry while, absent entry, there is no incentive to develop the technology.

Because of the diversity of the user community, it may be hard for them to solve the hold-up problem. Of course, this depends in the final analysis on how expensive the technology is compared to the benefit it brings to individual users. Clearly, it brings no benefit at all if there is no underlying choice of platforms. But even if there is such choice, its adoption still faces a prisoners' dilemma because the value of the technology is proportional to the degree of liquidity on the alternative platform, which in turn depends on the installed base of the technology. Collectively, there is an incentive to adopt it, but individually there is none. Discounting this, technology providers will not invest in developing it (it is, for various reasons, not a trivial matter simply to redeploy the technology developed for the United States).

This raises the question of whether an entrant can solve the problem by, for example, integrating with a technology provider or otherwise subsidizing the development of the technology.

3. Dominance in Trading Service Markets

As we have considered this issue elsewhere,[17] we briefly state our hypotheses here. Incumbent exchanges currently may possess dominant positions over trading services in the instruments that they trade. This possibility arises notably because of

17 S. Greenaway, *Competition between Stock Exchanges: findings from DG Competition's investigation into trading in Dutch equities*, 3 COMPETITION POL'Y NEWSL. 69-71 (Autumn 2005).

the exclusion of off-exchange alternatives from the relevant market due to the fact that, at least in the case of Dutch equity trading that the Commission investigated, they did not appear to constitute a significant competitive restraint. Similar conclusions on MTFs were reached by the U.K. Competition Commission in their investigation of possible bids for the London Stock Exchange.[18]

4. Incentives to Enter and Entry Inefficiencies

Even if barriers to entry can be overcome, there are still a couple of issues relating to potential competition between exchanges.

First, there may be strategic reasons to avoid entry. These arise because of the characteristics of the market post-entry and, in the case of an incumbent in another geography, the risk of retaliatory entry in the home market.

On the assumption that trading is largely a commodity business, it is expected that comparable players (such as two exchanges) would compete ex-post in prices, resulting in the Bertrand outcome, namely competing away of profit. A potential entrant might discount this outcome and realize that profits in the market post-entry will be insufficient to recuperate the costs of entry, even if these are quite low. Therefore, only the prospect of ejecting the incumbent entirely will induce entry.

INCUMBENT EXCHANGES CURRENTLY MAY POSSESS DOMINANT POSITIONS OVER TRADING SERVICES IN THE INSTRUMENTS THAT THEY TRADE. THIS POSSIBILITY ARISES NOTABLY BECAUSE OF THE EXCLUSION OF OFF-EXCHANGE ALTERNATIVES FROM THE RELEVANT MARKET DUE TO THE FACT THAT THEY DO NOT APPEAR TO CONSTITUTE A SIGNIFICANT COMPETITIVE RESTRAINT.

In this respect, the level of pre-entry rents being earned by the incumbent does not matter. Their level will neither induce entry, nor, by corollary, will the threat of entry discipline pricing. The reason for this is that, in a classical industry, capacity constraints limit the pricing response of an incumbent monopolist, so that post-entry, rents persist for at least a time even if goods are undifferentiated. In exchange trading, prices can be bid down to zero after entry because of the fear of massive liquidity shifts over a short space of time.

Even if the incumbent might be ejected, the prize might still not be worth the cost. If the challenger's home market is contestable, an incumbent facing price competition has an incentive to enter that market if it believes it has a lower overall level of fixed costs. If it fails to do so, the challenger can win the market even if it is less efficient, because it can fully subsidize its activity in the new market on the profits of the old. Anticipating retaliatory entry, a potential challenger may not enter, particularly if it does not believe it is intrinsically more effi-

18 *See Competition Commission – Inquiry – London Stock Exchange plc*, *at* http://www.competition-commission.org.uk/inquiries/ref2005/lse/index.htm (last visited Feb. 9, 2006).

cient. Even if it does believe that it could win a war of attrition, victory is likely to come with a hefty price tag in lost earnings, as the wounded incumbent would continue to cover its variable costs even at a very low price level and, thus, would not immediately exit. Compared to this scenario, a merger or takeover—if it wins regulatory approval—is likely to be less painful.

A corollary to this is what happens when a challenger enters on the basis of a home-market position that is not contestable. In this case, it can behave very aggressively. It will be unexposed to losses in the contested market and, regardless of whether it is more efficient or not, it can effectively bleed the incumbent until it withdraws. Since it has no need or use for the incumbent's assets—all it wants is its market position—it may achieve in this way the benefits of a takeover without paying anything at all. This outcome improves static welfare relative to two segmented markets and would have dynamic benefits as well, but if the challenger is not more efficient than the incumbent, then this outcome is not the best that can be achieved. Then the question that arises is whether the aggressive pricing behavior of the challenger is compatible with competition law. While this behavior looks predatory, it might very well not require pricing below any traditional cost standard, and so an assessment of its compatibility would necessarily lead into new territory. This would merit further debate.

B. COMPETITION IN CLEARING

In respect of clearing, two models are sometimes advanced. One is competition for the market, the other competition in the market. The latter implies interoperability between CCPs. We offer a couple of further thoughts.

First, with regards to competition for the market, this model seems difficult to apply because there is no public authority to organize the competition. Where there are instances of such competition, the possibility cannot be ignored that, at least to some extent, it might have helped transfer monopoly rents from clearing to trading. As just demonstrated, the power of trading over clearing can be used to prevent competition in trading, or at least result in no such competition occurring in practice. Further reinforcing these powers, therefore, may be unwise. Moreover, any such competition would not necessarily have anything to do with the service provided to users. It is not evident that, under these circumstances, investment by the incumbent clearer would occur when it is socially optimal to do so. The need to write assets off over a shorter time span might result in higher prices. Also, the exchange may have an incentive to change clearer when users have none, imposing costs on them. Obviously, it is even less realistic to expect competition to occur in a vertical silo. Therefore, current formulations of possible competition for the market appear naïve.

Interoperability of CCPs could be achieved and would bring benefits, although it would seem to only amount to competition in the market to a limited degree. The major benefit of achieving interoperability of CCPs would be that it would allow non-clearing members of foreign exchanges to clear their trades through their customary CCP rather than through a foreign agent bank. By clearing more trades through a single CCP, average collateral costs would fall at the same time as the foreign agency costs would be avoided. Having this option available might also induce more trading on foreign platforms, thereby deepening liquidity.

The costs of this solution would be considerably lower than sometimes alleged. Since any single user would have an account at a single CCP, all its exposure would be in relation to that CCP and the suggested solution would not increase collateral costs, even absent additional trading within CCP scope. It would not decrease netting efficiency either, because the CCPs could net their residual positions against each other. Doubtless, this solution requires technical standards and appropriate oversight to avoid moral hazard when a single CCP does not internalize the full risk of dealing with its members. Intuitively, however, this should be possible, although, as we have argued, it is unlikely to emerge simply under market conditions.

Interoperability of CCPs, of course, will only be of value if exchanges are required to route trades to the CCP of the user's choice. When the CCP is independent of the exchange and exchanges are unable to contractually foreclose competition in trading by the means described, greater competition at the clearing level should allow exchanges to capture greater rents than they otherwise would and so they may have an incentive to facilitate CCP interoperability. They would also internalize the benefit of wider access to their platform and the virtuous circle of increased liquidity. A vertical silo would probably not have this incentive because it is already able to capture all rents in its domestic market, and so it might be necessary to regulate in order to achieve full interoperability.

As already discussed, the problem of duplicating infrastructure costs does not arise because these are already sunk due to the current fragmentation of markets. A process leading to interoperability may, of course, eventually lead to full consolidation. It must be stated, though, that consolidation is not an alternative to interoperability since even already consolidated entities such as LCH.Clearnet and Euroclear still have not achieved full interoperability among their constituent historical components. Arguably, consolidation may make the path to interoperability smoother.

It remains to be discussed whether this scenario is real competition. As is apparent in the case of Virt-X—which does offer two CCPs—the choice of CCP by any given member is largely predictable. Thus, even under interoperability, we would expect bilateral monopoly largely to prevail. This may, however, not be the case for the largest players who presently use more than one CCP and could select any of them as their home CCP under interoperability. These players would achieve a significant advantage under this scenario.

C. COMPETITION IN SETTLEMENT

Turning to settlement, the option to internalize settlement or make use of an agent does not constitute a net competitive constraint on the pricing of a CSD. The same is true for settlement netting, except insofar as this may influence the share of the monopoly profit that the CCP is able to obtain and, hence, the share that remains for the CSD.

As we prove in the Annex (Section VII of this paper), it can be shown that CCP netting does not affect the CSD's profits. This result has a powerful corollary, since it implies that the presence or absence of a CCP does not influence the profit of the CSD. It is, in other words, irrelevant how many transactions the CSD has to process—settlement efficiencies do not influence its equilibrium level of profit. Note that it does still need to know the netting efficiency to set its price—this variable has not become irrelevant to its decision—but its profits remain unaltered.

THIS RESULT HAS A POWERFUL COROLLARY, SINCE IT IMPLIES THAT THE PRESENCE OR ABSENCE OF A CCP DOES NOT INFLUENCE THE PROFIT OF THE CSD.

This result gives a taste for the intuition that also lies behind the conclusion on settlement agency. However, the inability of settlement agency to disrupt the profit level of the CSD (if it is profit-maximizing) should not be taken to mean that the option to settle through an agent has no value at all. Settlement agents provide value-added services in respect of securities custody and, because of their ability to offer a broader bundled offering, may be convenient in other respects. In this paper, we have not explored competition in custody and any conclusions in this respect would be premature.[19]

VI. Conclusions

In discussing trading and post-trading markets, there is sometimes a tendency to generate theoretical solutions that take no account of inherited structures. This is, at best, unrealistic. When we take the existing landscape as our starting point and consider how it can be pragmatically improved, the role of competition in generating incremental efficiencies may take on greater importance than in a world in which economic analysis can leapfrog to theoretically optimal market structures. This would then justify increased attention on the part of competition authorities to the problems that the sector poses, even if they may appear complex. This, of course, is certainly not to say that regulation is unimportant—it may even be critical.

19 *See, e.g.*, J.-C. Rochet, The Welfare Effects of Vertical Integration in the Securities Clearing and Settlement Industry, (preliminary draft, IDEI, University of Toulouse) (Jun. 2005), *available at* http://www.cass.city.ac.uk/conferences/clearingandsettlement/Rochet.pdf.

In looking at current industry structures and incentives to achieve greater efficiency, there are both a number of possible market failures that may necessitate regulation, and a number of instances in which market players appear to have an incentive to adopt arrangements that might be considered under competition law. Whether the observed limited degree of competition, consolidation, and interoperability can be ascribed to such behavior and such market failures, is clearly more of a moot point. Given the macroeconomic issues at stake though, monitoring of the sector by competition authorities is a necessary accompaniment of regulatory efforts to achieve closer integration.

VII. Annex: CCP Netting and CSD Profits

For this analysis, assume a world in which the price charged for trading by the exchange and by the CCP is set first and the CSD then responds. (This analysis ignores fixed and ad valorem fees and assumes that each layer sets only a per-transaction fee. It also ignores exchange fees for unmatched trades.) There are n similar exchange members, all of which are also members of the CCP and CSD. They trade a single share between each other and their individual demand for trading in one clearing cycle is Poisson-distributed with mean equal to $M - p$, where M is a constant, $p = p_{TC} + p^*_S$ is the total price charged by all three layers, and p^*_S is the expected effective unit settlement price (i.e., the net settlement price divided by the number of transactions that are netted). Trading is possible only in single blocks of a given size. As a result, in each clearing cycle, there are on average $(M - p)n$ transactions, resulting in cn net positions where the CCP is creditor or debitor corresponding to each of the members with $c(p_S) <= 1$. These net positions are then forwarded to the CSD for settlement. Knowing the realization of the trading and clearing prices and of the residual demand function, the CSD then sets its price to maximize profits.

On the realistic assumption that the CSD has negligible variable costs, its objective function is simply to set p_S such that it maximizes $E(p_S cn)$, which is obviously the same value as maximizes $E(p_S c)$ given that n is known and positive. c, however, is the probability that any given member trades at least once. Given the Poisson probability distribution, it can be shown that it is equal to $1\text{-}e^{p-M} = e^{p_{TC}+p^*_S-M}$. Since $p^*_S = p_S/(M - p)$, and setting $A = M - p_{TC}$, then $p^*_S = p_S/(A - p^*_S)$. Since this is recursive, the math quickly becomes complex. However, it is not necessary to solve the first order condition, since it is enough to show that, when certain characteristics of the model are changed, the maximization problem faced by the CSD does not change.

Now assume there is lesser efficiency of netting in the CCP. In this case, instead of passing through cn settlement instructions to the CSD, it passes through a multiple of this, say kcn where $k > 1$. (For the proof, it is not important to consider why, if at all, such inefficiency could arise in practice, but it could be, for instance, the case if the CCP served two platforms and there was no fungibility—although in this case c would also change). In this case, the CSD will collect kcp'_S per member in revenue and this is the amount it will try to maximize. It is clear that it chooses $p'_S = p_S/k$ because, in this case, the expected average unit cost of settlement is unchanged and it is this that determines c. Total revenue per user is then $kp'_S = p_S$, which is divided over the total number of trades per user, $(M - p)c$, to give exactly the same formula for p^*_S as before.

Volume 2 | Number 1 | Spring 2006

Competition Policy International

Reply to Winter's "Vertical Restraints and Antitrust Policy: A Reaction to Cooper, Froeb, O'Brien, and Vita"

James Cooper, Luke Froeb, Daniel O'Brien, and Michael Vita

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1. *Competition Policy International* is a free publication. To order or download additional copies, visit eSapience.org.

Reply to Winter's "Vertical Restraints and Antitrust Policy: A Reaction to Cooper, Froeb, O'Brien, and Vita"

James Cooper, Luke Froeb, Daniel O'Brien, and Michael Vita

In the Autumn 2005 issue of *Competition Policy International*, we published an article on the antitrust policy implications of the theoretical and empirical literature on vertical restraints.[1] In an accompanying comment,[2] Professor Ralph Winter claims that we are advocating an enforcement standard that in any particular case would ignore case-specific evidence of the restraint's effects. He also claims that we commit an "analytical error"[3] in our discussion of how distortions in promotional incentives may motivate the use of resale price maintenance (RPM).

1 J. Cooper, L. Froeb, D. O'Brien, & M. Vita, *Vertical Restrictions and Antitrust Policy: What About the Evidence*, 1(2) COMPETITION POL'Y INT'L 45–63 (2005).

2 R. Winter, *Vertical Restraints and Antitrust Policy: A Reaction to Cooper, Froeb, O'Brien, and Vita,* 1(2) COMPETITION POL'Y INT'L 75–88 (2005). A comment was also published in the same issue by F.M Scherer (*Comment on Cooper et al.'s Vertical Restrictions and Antitrust Policy,* 1(2) COMPETITION POL'Y INT'L 65–74 (2005). We respond to many of the points raised by Professor Scherer *in* J. Cooper, L. Froeb, D. O'Brien & M. Vita, *Vertical Antitrust Policy as a Problem of Inference*, U.S. Federal Trade Commission (Feb. 18, 2005), *available at* http://www.ftc.gov/speeches/froeb/050218verticalecon_respcrit.pdf.

3 Winter, *supra* note 2, at 82.

James Cooper is Assistant Director, Office of Policy Planning, Federal Trade Commission. Luke Froeb is William Oehmig Associate Professor at Owen Graduate School of Management at Vanderbilt University. Daniel O'Brien is Chief, Economic Regulatory Section, in the Economics Analysis Group of the Antitrust Division, U.S. Department of Justice. Michael Vita is Assistant Director for Antitrust, Bureau of Economics, Federal Trade Commission. The views expressed herein are the authors' own and do not purport to represent the views of the U.S. Department of Justice, the Federal Trade Commission, or any Commissioner.

Neither criticism is accurate. We agree with Professor Winter that individual cases should be judged "on their own merits." We do argue, however, that the plaintiff's burden in vertical restraints case should be high, and we place significant weight on both historical and case-specific empirical evidence.[4]

His claim that we commit an analytical error in our discussion of promotional incentives also is incorrect. We clarify this point below.

I. On Using Case-Specific Evidence in Antitrust Investigations

Our article observed that economic theory provides ambiguous predictions about the welfare effects of vertical restraints. Moreover, we explained that when practitioners attempt to use theory to help discern the effects of vertical restraints in any particular case, they confront a difficult inferential problem:

> "Not only must they decide which model best applies to the particular factual circumstances in which the restraint has been adopted, they also must then determine whether the model chosen has the particular combination of parameters that would result in an anticompetitive equilibrium."[5]

We pointed out that this is a difficult exercise, given the various uncertainties that are involved, and concluded that the ability of theory to guide practitioners is "quite limited,"[6] and that enforcement decisions necessarily will be subject to "substantial uncertainty."[7] Nowhere in this discussion did we say that, in any

4 Remarkably, in earlier academic publications, Professor Winter advocated an explicit policy of per se legality for vertical restraints:

> With a single firm upstream, and therefore no possibility of beneficial effects on interbrand competition, we find that restraints nonetheless generally increase market surplus in our framework. Although examples can be constructed in which welfare decreases with restraints, a rule superior to per se legality of purely vertical restraints has not, in our view, been offered.

See F. Mathewson & R. Winter, *On Vertical Restraints and the Law: A Reply*, 19(2) RAND J. ECON. 300 (1988).

5 Cooper et al., *supra* note 1, at 45.

6 *Id.* at 47.

7 *Id.* at 48.

given instance, the competitive implications of vertical restraints should not be assessed on their own merits. We did say that doing so is typically "difficult"—a point that should be uncontroversial.

Given the significant uncertainty that characterizes the equilibrium consequences of vertical restraints in any particular application, our article went on to discuss how an optimal policy toward vertical restraints "could be modeled as a process whereby decision makers use observed data to update their prior beliefs about the likely efficiency of a given vertical restraint, yielding a posterior belief."[8] We explained that a survey of the empirical literature on vertical restraints shows that, in the cases that have been studied to date, vertical restraints almost always have been found to be pro-competitive or competitively neutral. This suggests that an optimal decision process would start with strong priors that vertical restraints are pro-competitive.[9] With respect to the evidence in a particular case, we noted, "if empirical evidence is difficult to interpret, these observations will cause little, if any, modification to these prior beliefs."[10] This hardly is equivalent to "forcing appropriate antitrust decisions to rely on prior empirical evidence rather than case-specific facts,"[11] as Professor Winter contends.

Finally, we discussed the implications of our analysis for antitrust policy. We noted that "to the extent that theory provides little guidance in classifying evidence beyond allowing us to determine safe harbors, a decision maker's beliefs that a specific vertical practice is pro- or anticompetitive should closely mimic his or her prior beliefs regarding such practices in general."[12] This statement simply is a logical conclusion about the weight that murky evidence should receive when a rational decision maker optimally updates his prior beliefs. It is not a call for case-specific evidence to be ignored. Indeed, in the conclusion, we suggested a case-specific approach in which policymakers "draw inferences about the competitive effects of the restraint by comparing markets with and without the restraint to determine the effect of the restraint."[13] We noted that this approach could involve a comparison of the same market before and after adoption of the constraint, or a comparison of a cross-section of markets in different geographic areas. Obviously, the quality of the experiment, and how closely it mimics the

8 *Id*.

9 Of course, the strength of these priors may vary by the type of restraint at issue because not all vertical practices have been subject to the same level of empirical research.

10 Cooper et al., *supra* note 1, at 48.

11 Winter, *supra* note 2, at 87.

12 Cooper et al., *supra* note 1, at 61.

13 *Id.* at 63.

effect of the restraint, will affect the weight that a decision maker should attach to such evidence.

II. Professor Winter's Critique of Our Discussion of RPM

In our discussion of the various possible motivations for vertical restraints, we observed that, in numerous models, RPM induces retailers to spend more on promotion than they would without RPM. In his comment, Professor Winter claims that our discussion of this issue commits "an analytical error."[14] We stated, "[i]f the manufacturer's profit margin for additional sales is large in relation to the retailer's . . . the retailer rationally will provide a lower level of promotion than is optimal for the manufacturer."[15]

Referencing this passage, Professor Winter states, "the authors err . . . in stating that inadequate retailer incentives can be traced to differences between the wholesale margin and the retail margin."[16] Professor Winter's criticism, however, ascribes greater precision to our statement than is warranted. Note that our statement does not say how large the manufacturer's margin must be relative to the retailer's margin for the retailer to choose less promotion than is optimal for the manufacturer. The statement, therefore, is quite weak and is true across a wide range of imperfectly competitive environments. It holds, for example, in the environment examined in Mathewson and Winter's seminal analysis (1984) of vertical restraints by a manufacturer that charges two-part tariffs to imperfectly competitive retailers engaged in promotion that may or may not spillover to their rivals.[17] It also holds in the same environment when firms restrict attention to linear input prices.[18] One can imagine theoretical environments in which our

14 Winter, *supra* note 2, at 82.

15 Cooper et al., *supra* note 1, at 49.

16 Winter, *supra* note 2, at 84.

17 *See* F. Mathewson & R. Winter, *An Economic Theory of Vertical Restraints*, 15(1) Rand J. Econ. 27-38 (1984). Since the manufacturer can employ two-part tariff contracts in this model, its optimal level of promotion is the same as that of a fully integrated firm. Absent vertical restraints, it turns out that any positive margin for the manufacturer is large enough relative to the retail margin to cause retailers to choose less promotion than is optimal for the manufacturer.

18 Under linear input pricing, if the retailer's margin is sufficiently small relative to the manufacturer's margin at the chosen wholesale price, the retailer will select a level of promotion lower than the manufacturer would choose if the latter could establish the level of promotion at the time the wholesale price is set. If the manufacturer and retailer cannot agree on the level of retail promotion in advance (an assumption required for retail promotional incentives to motivate RPM in the first place), then they may turn to resale price maintenance to improve the retailer's incentives.

statement is too strong.[19] However, our statement is weak enough to cover most cases of interest. More precise statements about the relationship between margins and promotional incentives would require detailed assumptions, which would go well beyond the appropriate scope of a policy journal.

III. Conclusion

We do not advocate a standard of per se legality for vertical restraints. We believe that case-specific evidence is relevant to determining the legality of any particular use of a vertical restraint, and our article suggested a number of ways that such evidence could be brought to bear by antitrust investigators. That said, we defend unapologetically a vertical restraints enforcement standard that forces plaintiffs to bear a high burden of proof. Theory tells us that vertical restraints can be good or bad, but the weight of the best available empirical evidence comes down overwhelmingly on the "good" side of the scale.

19 For example, in a private conversation, Professor Winter pointed out that an exception occurs when perfectly competitive retailers sell to customers with identical preferences over quality. In that special case, retailers choose the manufacturer's preferred level of promotion even when their margins are zero.

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Rejoinder to Cooper, Froeb, O'Brien, and Vita's Reply

Ralph A. Winter

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Rejoinder to Cooper, Froeb, O'Brien, and Vita's Reply

Ralph A. Winter

In this rejoinder, I first respond to the discussion in Cooper, Froeb, O'Brien, and Vita's "Reply to Winter" of a technical point, the relationship between retailer incentives and retailer margins, and then set out our common ground and remaining differences on the broader theme of theory and evidence in vertical restraints cases.[1]

Cooper et al. stated in their original article that a retailer will provide a lower level of effort than is optimal for the manufacturer when the retailer's margin is small relative to the manufacturer's margin. I claimed in my comment on the article that low retail margins do not necessarily lead to inadequate retailer incentives for promotion. One counterexample to Cooper et al.'s general claim is the framework developed in my comment in which a manufacturer and its retailers strike contracts that maximize their joint wealth. Another counterexample is the simple benchmark in which the sales-generating effort by retailers is in effect adding to product quality, in which all consumers have identical preferences for quality, and in which the market structure consists of a monopoly upstream and perfect competition downstream. (Retailers in this setting produce exactly the quality that is optimal for the manufacturer, yet the retailer margin

1 This article is a rejoinder to J. Cooper, L. Froeb, D. O'Brien, & M. Vita, *Reply to Winter's "Vertical Restraints and Antitrust Policy: A Reaction to Cooper, Froeb, O'Brien, and Vita,"* 2(1) Competition Pol'y Int'l 189-193 (Spring 2006). For the original articles, *see* J. Cooper, L. Froeb, D. O'Brien, & M. Vita, *Vertical Restrictions and Antitrust Policy: What About the Evidence*, 1(2) Competition Pol'y Int'l 45–63 (Autumn 2005) and R. Winter, *Vertical Restraints and Antitrust Policy: A Reaction to Cooper, Froeb, O'Brien, and Vita,* 1(2) Competition Pol'y Int'l 75–88 (Autumn 2005).

The author is Professor and Canada Research Chair in Business Economics and Public Policy at Sauder School of Business, University of British Columbia.

is zero.) In their reply, however, Cooper et al. provide examples of assumptions under which their claim holds.[2] Low retailer margins may or may not be associated with inadequate retailer incentives. Any attempt to be more precise about the relationship between the retailer margin and the distortion in retailer incentives is perhaps unproductive since both variables are endogenous.

Our differences on this technical detail should not distract the reader from the interesting contributions of Cooper et al. (2005). The most striking of their themes, as summarized in the introduction to the article, is the following statement: "We argue that economic theory actually provides policymakers with very little guidance as to whether vertical restraints are likely to be beneficial or harmful in any particular factual setting."[3]

This argument is highly provocative to an economist who believes that theory is not just valuable, but essential, in interpreting the facts of a case. Cooper et al. do not suggest that case evidence and economic theory are never reliable of course, but their strong emphasis is on prior evidence on the relative frequency of pro-competitive versus anticompetitive effects of the particular restraint at issue in a case. I defend the opposite position in my comment: the heart of a vertical restraints case is the application of theory to the factual setting of the case. This factual setting includes conventional evidence on market definition and indicators of market power, but what is especially important is case evidence that allows testing of pro-competitive versus anticompetitive theories of the restraint at issue. Some case evidence, in "naked exclusion" cases for example, is consistent with anticompetitive theories. Other case evidence suggests pro-competitive theories are at work.

I exaggerate the difference in our views, however. As in competition policy generally, prior evidence of the type that Cooper et al. emphasize is vital. Its role is in determining where the burden of proof in the court's assessment of case evidence should lie and in this sense the two kinds of evidence are complementary. In merger analysis, to take an example outside the vertical restraints context, the burden of proof in demonstrating a lessening of competition lies with the plaintiff or government because the overwhelming majority of mergers are pro-competitive. In price-fixing cases, it does not.

Cooper et al.'s review of the evidence on vertical restraints provides support for the position that a strong burden of proof in vertical restraint cases should lie on the side of government intervention, not just on some vertical restraints as it

2 My description of Cooper et al.'s statement as an "analytical error" was inappropriate. I would have more reasonably indicated that the conditions under which the statement is true are not clearly stated in the article and may well be violated in practice.

3 Cooper et al. (2005), *supra* note 1, at 47.

has since *Sylvania*[4] in 1977 and *State Oil v. Khan*[5] in 1997, but also for minimum resale price maintenance—the most common vertical restraint (when it was legal).[6] A sensible policy, in my view, would allow restraints with purely vertical effects. It is theoretically possible that a monopolist would use resale price maintenance to shift the mix of retail price and service competition to its own benefit and to the detriment of consumers, but (consistent with Cooper et al.'s message) economic theory does not offer a clear delineation of when a purely vertical use of restraints would be anticompetitive. In many cases, the purely vertical use clearly increases welfare.

On the other hand, case studies, which must be considered the essential component of prior empirical evidence, reveal instances where vertical restraints are harmful. Economic theory does, in fact, provide clear guidance as to the impact of restraints with horizontal effects that are exclusionary at the product level (*Nielsen*), exclusionary at the retail level (*Toys "R" Us*), or collusive (*General Electric*).[7] A sensible rule of reason would prohibit restraints that have detrimental horizontal effects, with the burden of proof falling on the government or plaintiff. The details of an optimal rule of reason on vertical restraints would be a challenge, since the dividing line between purely vertical and horizontal effects can be elusive and, as in any area of competition policy, there will be cases on the margin. Economic theory and prior empirical evidence would provide the foundation for designing the rule. Economic theory and case evidence would be the key instruments in applying the rule.

4 Cont'l T.V. Inc,. v GTE Sylvania Inc., 433 U.S. 36 (1977), in which the U.S. Supreme Court overruled U.S. v. Schwinn, 388 U.S. 365 (1967), as Cooper et al. discuss, and held that non-price vertical restraints were to be judged under a rule of reason, with the burden of proof of adverse competitive effects falling on the plaintiff.

5 State Oil Company v. Barkat U. Khan, U.S. 118 S.Ct. 275 (1997), in which the U.S. Supreme Court overruled Albrecht v. Herald Co., 390 U.S. 145 (1968), and decided the vertical maximum price restraints should be judged under a rule of reason.

6 I have previously advocated a policy of per se legality towards restraints that are purely vertical, such as the use of resale price maintenance and territorial restrictions by a monopolist. Cooper et al. note this in footnote 4 of their reply. I have not advocated per se legality of vertical restraints where horizontal effects are involved, as is the case with naked exclusionary restraints or where resale price maintenance facilitates collusive pricing.

7 *Nielsen* (Canada (Director of Investigation and Research) v. The D & B Companies of Canada Ltd. (1995), 64 C.P.R.3d 216 (Comp.Trib. 1995)) is discussed in my comment. Toys "R" Us v. FTC, 221 F.3d 928, 937 (7th Cir. 2000) and United States v. General Electric Co., 358 F. Supp. 731 (1973) are discussed in F.M. Scherer's comment on Cooper et al. (2005).

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Court of First Instance Upholds Prohibition of *General Electric/Honeywell*

Shaun Goodman

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Court of First Instance Upholds Prohibition of *General Electric/ Honeywell*

Shaun Goodman

On December 14, 2005,[1] the European Court of First Instance (CFI) upheld the European Commission's 2001 prohibition of a proposed merger between General Electric (GE) and Honeywell (the "Decision").[2] The Decision's partial reliance on conglomerate effects theories had been controversial at the time, and the Commission was criticised in strong terms by U.S. regulators that had approved the transaction. Following a recent series of judicial reversals of EC merger prohibition decisions,[3] the CFI's confirmation of the Decision must have come as a relief to the Commission. However, the grounds on which the Decision was upheld were narrow and, in respect of the most controversial aspects of the Decision—namely its reliance on alleged vertical and conglomerate effects—the CFI found that the Commission had committed manifest errors. Together with recent judgments of the CFI and European Court of Justice (ECJ) in *Tetra Laval*, the CFI's judgment in *GE/Honeywell* confirms the high standard that the Commission must meet to prohibit a conglomerate merger, thereby making it less likely that such transactions will be prohibited in the future.

1 Case T-210/01, General Electric v. Commission, 2001 O.J. (C 331) and Case T-209/01, Honeywell v. Commission, 2001 O.J. (C 331).

2 Commission Decision 2004/134/EC, General Electric/Honeywell, 2004 O.J. (L 48) 1 [hereinafter *GE/Honeywell*].

3 Case T-342/99, Airtours plc v. Commission, 2002 E.C.R. II-3585; Case T-310/01, Schneider Electric v. Commission, 2002 E.C.R. II-4071; and Case T-5/02, Tetra Laval v. Commission, 2002 E.C.R. II-4381 [hereinafter *Tetra Laval*].

Shaun Goodman is a partner in the London office of Cleary Gottlieb Steen & Hamilton LLP.

I. The Decision

On July 3, 2001, following an in-depth investigation, the Commission prohibited GE's proposed merger with Honeywell. The Commission identified the following four main competition concerns.

1. First, the Commission determined that GE held a pre-existing dominant position in large regional jet engines, which would be strengthened by the addition of Honeywell's competing business on this market (i.e., horizontal effects).

2. Second, the Commission found that the combination of GE's and Honeywell's activities in the markets for corporate jet engines and small marine gas turbines would create dominant positions (i.e., horizontal effects).

3. Third, the Commission found that Honeywell had a strong position in engine starters (which are a necessary input for creating a full engine package). The Commission considered that GE's acquisition of Honeywell's engine-starter business would strengthen GE's pre-existing dominance in large commercial jet engines because it would allow GE to disrupt supplies of Honeywell engine starters to GE's engine competitors (i.e., vertical effects).

4. Fourth, and most controversially, the Commission concluded that the combination of GE's dominant position in large commercial jet engines and Honeywell's leading positions in a broad range of avionics and non-avionics systems would create a dominant position in the avionics markets through two types of conglomerate effects:

    - The first effect would arise from GE's reliance on its leasing subsidiary GE Commercial Aviation Service (GECAS), which buys aircraft from manufacturers and leases them to airlines. The Commission held that GE could use GECAS as a commercial lever by offering airframe manufacturers and airlines concessions in return for specifying Honeywell products on the aircraft they purchase. The Commission found that GE had used GECAS in a similar way to promote its large commercial jet engines.

    - The second effect would arise from GE's bundling of its large commercial jet engines with Honeywell's avionics products. According to the Commission, such bundling could take the form of pure bundling (i.e., refusing to make available the engines without the avionics), technical bundling (i.e., integrating the engines and the avionic systems) or mixed bundling (i.e., offering a discount if customers take both the engines and the avionics from GE). As a result, GE could extend its dominance from large commercial jet engines to avionics. Conversely, GE's bundling of avionics systems in which Honeywell held a leading position would also reinforce its pre-existing dominance in large commercial jet engines.

II. The Judgment

The CFI subjected each of the Commission's findings to close examination. It referred to the judgment of the ECJ in *Tetra Laval*, which had recognized that while the Commission enjoys a margin of discretion in "appraisals of an economic nature," the CFI was obliged to review whether the Commission's "evidence [. . .] is factually accurate, reliable, and consistent, [ ] whether that evidence contains all the information which must be taken into account in order to assess a complex situation and whether it is capable of substantiating the conclusions drawn from it."

The CFI held that the Commission's findings on the strengthening of a dominant position in large commercial jet engines and the creation of a dominant position in avionics were not sufficiently supported. At the same time, however, it confirmed the Commission's findings on the strengthening of a dominant position in large regional jet engines and the creation of dominance in corporate jet engines and small marine gas turbines. The CFI concluded that the findings on large regional jet engines, corporate jet engines, and small marine gas turbines were each sufficient to support a prohibition of the proposed merger, and that the Commission's errors as regards large commercial jet engines and avionics, therefore, did not justify annulment of the Decision.

A. STRENGTHENING OF GE'S DOMINANT POSITION IN LARGE REGIONAL JET ENGINES THROUGH HORIZONTAL OVERLAPS

The CFI confirmed that the Commission had properly established GE's existing dominance on the basis of market share data. GE's engines accounted for 60-70 percent of large regional aircraft still in production and for 90-100 percent of order backlogs for large regional aircraft not yet in service. Third-party suppliers, other than Honeywell, were not active on the market at the time of the Decision.

The CFI confirmed the Decision's finding that GE's dominant position in large regional jet engines would have been strengthened notwithstanding the absence of direct competition between GE's and Honeywell's engines. At the level of airlines, there was no direct competition since airframe manufacturers only certified one engine for a given airframe, while, at the level of airframe manufacturers, there was no direct competition because GE's engines could only be used on a two-engine platform while Honeywell's engines could be used only on a four-engine platform.

The CFI upheld the Commission's determination that GE's and Honeywell's engines competed indirectly through the selection by airlines of complete aircraft equipped with different engines. Among other things, the CFI pointed to internal GE documents demonstrating that GE granted discounts on its engines in order to boost the sale of aircraft equipped with its engines. Accordingly, so the CFI reasoned, the merger would have eliminated existing competition between GE and Honeywell engines.

Given that the merger would have effectively given GE a monopoly in engines for large regional aircraft, the CFI rejected GE's contention that the acquisition of Honeywell would have had only a marginal impact on its position. The CFI noted that, under Article 82 of the EC Treaty, the greater the dominance of an undertaking, the greater its obligation to abstain from any conduct that is liable to weaken existing competition. By analogy, the CFI held, a company with a strong dominant position cannot contend that the acquisition of a competitor does not raise concern simply because that rival is already weak or merely exercises an indirect competitive constraint. Rather, the reduction of any residual competition is particularly harmful.

By analogy, the CFI held, a company with a strong dominant position cannot contend that the acquisition of a competitor does not raise concern simply because that rival is already weak or merely exercises an indirect competitive constraint.

Finally, the CFI found no fault with the Commission's rejection of commitments offered by GE in an effort to address this concern, since there were legitimate doubts as to whether the divestiture of Honeywell's large regional engine business would have created a viable business.

B. CREATION OF A DOMINANT POSITION IN CORPORATE JET ENGINES AND SMALL MARINE GAS TURBINES THROUGH HORIZONTAL OVERLAPS

The CFI confirmed the Commission's findings on the creation of a dominant position in engines for corporate jets and small marine gas turbines.

Corporate jet engines

The CFI agreed that the Commission could rely on the parties' market shares for its conclusion that the transaction would create a dominant position in corporate jet engines. The merged entity would have held 50-60 percent of the installed base of engines for corporate jets and 80-90 percent of engines for medium corporate jets. The CFI noted that, in line with past case law, such shares were in themselves sufficient to demonstrate dominance.[4]

Small marine gas turbines

The dispute on this point focused on the Commission's identification of a relevant market limited to gas turbines of 0-10 megawatts for marine applications. GE maintained that GE's and Honeywell's gas turbines did not compete with each other. The Commission's definition was based largely on responses to information requests received from three competing suppliers of small gas turbines. The CFI noted that one of the responses was ambiguous, one was consistent with the Commission's market definition, and one advocated a broader definition but

4 Case T-221/95, Endemol Entertainment v. Commission, 1999 E.C.R. II-1299; Case 62/86, AKZO v. Commission, 1991 E.C.R. I-3359; and Case 85/76 Hoffman-La-Roche v. Commission, 1979 E.C.R. 461.

confirmed that GE's and Honeywell's gas turbines competed with each other. In these circumstances, the CFI concluded that the Commission had not committed a manifest error in defining the relevant market.

The CFI considered it irrelevant that the Commission had failed to seek information from the only European customer of each of the parties since GE had not shown or alleged that this failure affected the Commission's finding. The CFI also considered it irrelevant that Honeywell's gas turbine had competed against GE's gas turbine in a bidding process only once during the last five years, since bids in maritime gas turbines were rare.

C. STRENGTHENING OF GE'S DOMINANT POSITION IN LARGE COMMERCIAL JET ENGINES THROUGH VERTICAL FORECLOSURE

1. GE's Existing Dominance

The CFI confirmed that the Commission was correct in finding that GE occupied, pre-merger, a position of dominance in the market for large regional jet engines. The CFI discussed three key issues:

- (i) whether the use of market shares as indicators of market power was appropriate in a bidding market;
- (ii) whether it was correct to attribute the market share of a 50/50 joint venture entirely to GE; and
- (iii) to what extent GE's reliance on GECAS strengthened its dominant position.

Market shares in bidding markets

The CFI agreed with GE that, in bidding markets where orders are large and infrequent, high market shares may not necessarily be indicative of dominance since shares may fluctuate significantly depending on recent wins and losses. At the same time, however, the CFI noted that GE had not only succeeded in maintaining its leading position over five years, but had also enjoyed the highest market share growth rate during this period. The CFI, therefore, concluded that the Commission could properly rely on GE's market shares for the assessment of its dominance. The CFI also observed that "lively competition on a particular market" does not rule out the existence of dominance.

Allocation of JV sales

The CFI confirmed that the Commission was correct in allocating the market share of CFM International (CFMI) (a 50/50 joint venture between GE and France's Snecma) entirely to GE, even though it rejected the Commission's suggestion that CFMI was a quasi-subsidiary of GE. The CFI highlighted the following elements:

- CFMI's engines did not compete with GE's engines. According to the CFI, GE and CFMI effectively acted as a single entity vis-à-vis competitors and customers.

- Snecma, unlike GE, did not and could not produce engines independently. Allocating part of CFMI's sales to Snecma, therefore, would not have reflected true market reality.

- GE's own annual reports attributed CFMI's market share entirely to GE.

Leveraging of GECAS

The CFI agreed that the Commission could treat GE's reliance on GECAS as an element that strengthened its dominance. GECAS had enabled GE to influence engine selection by airframe manufacturers and airlines and, thus, to win contracts that it would not have won through competition on the basis of price and technical quality alone. The effects of GECAS on engine competition differed depending on who selected the engine for a given aircraft type:

- If the airframe manufacturer selected the engine for a given aircraft type, GECAS's role as a large purchaser of aircraft would create a strong incentive for manufacturers to place GE engines on their new airframes, since GECAS had a well-established record of buying only GE-powered aircraft. GECAS accounted for 7-10 percent of all large commercial aircraft purchases. Aircraft manufacturers would know that if they did not select GE engines, GECAS would not purchase their airplanes and thus, they would be cut off from this portion of the market.

- If the airline selected the engine, GECAS as a leasing company could offer airlines concessions if they took GE's engines. In addition, GECAS could influence the choice of airlines indirectly by "seeding" the market with GE equipped aircrafts. Given the benefits available to airlines (e.g., in terms of lower maintenance costs) of using the same engine type across their entire fleet, GECAS's seeding policy created incentives for airlines to standardise their fleet on GE engines.

The Decision included evidence that GECAS had in fact played an important role in actual engine selection decisions by airframe manufacturers. In light of the Commission's evidence, the CFI rejected GE's objection that the Commission's economic theory (based on GECAS's relatively small share of total aircraft purchases) was "unorthodox." It was also irrelevant that the Commission had been unable to provide statistical data showing that GECAS actually had increased GE's overall market share. According to the CFI, the individual incidents described by the Commission were sufficient to demonstrate that GE had used GECAS to promote its engines and that this policy had met with success in individual cases. Moreover, GE's economists had not been able convincingly to show with their own statistical models that the use of GECAS had not had an impact on the market.

2. Strengthening of Dominance through Vertical Foreclosure

The CFI also followed most of the Commission's reasoning with respect to the risk of foreclosure arising from GE's acquisition of Honeywell's engine-starter business, but stopped short of endorsing the Commission's conclusions on this point. The CFI agreed that GE's engine competitors were dependent on Honeywell engine starters and that GE would have a commercial incentive to delay or disrupt supplies of Honeywell engine starters to its competitors post-merger. This was because engine-starter sales represented only a small fraction (around 0.2 percent) of the profits that GE could derive from additional engine sales. As a result, it would be to GE's advantage to forego profits from engine starters in order to win engine market share at the expense of its competitors.

The CFI rejected GE's objection that the Commission had not produced an "economic study" to prove GE's incentives and the likely market development. The CFI explained that where it is "obvious" that the merged entity will have the incentives to behave in a certain way, the Commission does not commit a manifest error in holding that it is likely that the merged entity effectively will behave in that way. In such circumstance, the "simple economic and commercial realities" of the case may constitute "convincing evidence" for supporting the Commission's conclusions, thus meeting the standard of proof set by the ECJ in *Tetra Laval*.

However, the CFI held that the Commission's analysis was incomplete because the Commission had failed to take into account the deterrent effect of Article 82. According to the CFI, a disruption of engine-starter supplies as contemplated by the Commission would "clearly amount to an abuse." The CFI pointed out that the abusive conduct need not take place in the market in which dominance is found to exist. It also noted that the more convincing the Commission's case on the effectiveness of the supply disruption would be, the more likely that the conduct would infringe Article 82. The need to consider the deterrent effect of Article 82 in assessing whether the strategic conduct was likely to take place was established by the CFI's judgment in *Tetra Laval* (which the Decision preceded). The CFI noted that while the Commission did not have to engage in an in-depth assessment of the deterrent effect of Article 82, it nevertheless required the Commission to undertake at least a "summary analysis based on the evidence available to it." Accordingly, the CFI concluded that the Commission had committed a manifest error of law by failing to discuss the possible deterrent effect of Article 82.

THE CFI POINTED OUT THAT THE ABUSIVE CONDUCT NEED NOT TAKE PLACE IN THE MARKET IN WHICH DOMINANCE IS FOUND TO EXIST.

D. CREATION OF A DOMINANT POSITION IN AVIONICS THROUGH CONGLOMERATE EFFECTS

As noted above, the Decision identified two types of conglomerate effects that the Commission alleged would create a dominant position in avionics:

(i) the leveraging of GECAS; and

(ii) the bundling of GE engines with Honeywell avionics.

The CFI found that the Commission's assessment was erroneous in both respects.

Leveraging of GECAS

Although, as noted above, the CFI endorsed the Commission's finding that GECAS played a role in establishing GE's pre-existing dominance in engines, the CFI held that the Commission had not provided sufficient evidence to demonstrate that GE would have extended the same practices to create dominant positions for Honeywell's avionic products. Such evidence could have consisted, for example, of an economic study of GE's incentives or documents demonstrating that GE effectively intended to use GECAS in favour of avionic products post-merger. It was not sufficient for the Commission simply to point to GE's reliance on GECAS for the promotion of engines and assume that the same mechanisms would apply with respect to avionics.

The CFI pointed out that GE's reliance on GECAS entailed costs in the form of the concessions that GECAS made to customers. In the case of engines, these costs were off-set by the revenue streams generated from after-sale services. Yet, in the case of Honeywell's avionics, the Commission had not examined whether the revenue generated from avionics sales would be capable of compensating the costs of relying on GECAS and, therefore, whether such reliance would be worthwhile for GE.

The CFI, moreover, found that the Commission had not proven that GE's reliance on GECAS for the promotion of avionic products would effectively lead to the creation of a dominant position. The Commission had ignored the fact that GECAS was only active in the area of large commercial and large regional aircraft, while Honeywell's avionic products were also sold for other aircraft. In addition, the Commission's analysis had failed to distinguish properly between the different avionics product markets. As a result, the Commission failed to demonstrate what the likely impact of the transaction would have been on each relevant market.

Bundling

Similarly, the CFI held that the Commission had not provided sufficient evidence to demonstrate that GE would have an incentive to engage in bundling of engines and avionic products. The CFI cited the following main factors in this respect:

- The scope for bundling of engines and avionic products was limited because the two products were not always selected by the same operators: engines might be selected by the airframe manufacturer while avionics might be selected by airlines, or vice versa.

- Bundling would entail costs, since GE would lose customers that preferred a different combination or would have to give customers discounts to overcome such preferences. Yet, the Commission had not analysed to what extent the increased sales of avionic products would off-set such costs. It was important for the CFI's assessment that while GE was dominant in large commercial jet engines it still faced viable competition in this area.

- Snecma, which jointly controlled GE's CMFI engines joint venture, would have no incentive to sacrifice part of its profits in order to promote Honeywell avionics through a bundling strategy.

- The Commission could not simply point to Honeywell's past practice of bundling different avionic products as evidence of likely future engine/avionics bundling, since the price of engines was markedly higher than the price of avionics. As a result, it was not excluded that the commercial dynamics of an engine/avionics bundle were different from bundling avionics.

- It was also not sufficient simply to refer to the "Cournot effect of bundling," which describes the advantages that companies can derive from a large range of products. As the economists of one of the Commission's supporters recognised in a newsletter that GE presented to the Court, the Cournot effect requires a detailed analysis of the necessary discounts and expected shifts in sales, which the Commission had not made. According to the CFI, both pure and mixed bundling would have infringed Article 82. Yet, the Commission failed to discuss the possible deterrent effect of Article 82 even in summary form.

III. Analysis

The *GE/Honeywell* judgment provides interesting insights and valuable clarifications in a number of areas, including on the Court's standard of review, the relevance of economic evidence, the relevance of market shares, theories of vertical and conglomerate effects, the assessment of horizontal overlaps, and the interplay of merger control rules and Article 82.

The CFI's standard of review

The judgment confirms that the CFI will review closely the evidentiary basis of Commission decisions. At the same time, however, the Court will exercise restraint in reviewing conclusions drawn by the Commission from that evidence. The CFI's willingness to grant the Commission a margin of discretion in respect of questions of a complex economic nature is illustrated by the CFI's discussion of GE's dominance, issues of market definition, the competitive interaction between GE and Honeywell, and the assessment of GE's commitments.

The CFI's reluctance to annul the Decision even where it has found substantial parts of that decision to be defective is also noteworthy. The CFI's "independent pillars" theory—which conforms to a long-standing practice of the ECJ—requires appellants to bring an effective challenge against all independent grounds of a decision. Thus, the CFI was able to dismiss Honeywell's parallel appeal in summary form because it had not challenged the Decision's findings in respect of all markets.

Economic evidence

The judgment confirms that the Commission enjoys a considerable degree of flexibility in the type of evidence that can be relied on to discharge its burden of proof. The Court did not require the Commission to support its conclusions with any specific type of evidence. The judgment instead suggests that the Commission may choose among various types of evidence, including economic studies, internal documents, concrete factual examples, or responses from market participants.

The judgment confirms that the Commission enjoys a considerable degree of flexibility in the type of evidence that can be relied on to discharge its burden of proof.

Conglomerate effects

The CFI's judgment does not exclude application of the Commission's conglomerate effects theory, although, in endorsing and applying the framework developed by the EC Courts in *Tetra Laval*, it confirms the high evidentiary standard that must be met by the Commission when it challenges transactions based on their conglomerate effects.

Consistent with the EC Courts' judgments in *Tetra Laval*, the CFI in *GE/Honeywell* required the Commission to demonstrate that the merged entity will be likely to engage in the conduct anticipated by the Commission. The judgment provides that the Commission must analyse the likelihood of the merged entity's future conduct on the basis of the entity's economic incentives and any factors that may deter it from adopting the conduct in question. In making its assessment, the Commission may rely either on internal documents or an examination of the parties' commercial interests in the relevant market at issue. The circumstance that one of the merging parties is engaging in similar conduct on a different market, on the other hand, will not generally be sufficient to support an adverse finding.

As regards the assessment of the anticipated competitive impact of the post-transaction conduct, the judgment's reasoning suggests that the CFI will generally grant the Commission a margin of discretion in this respect, provided the Commission clearly identifies the relevant markets that will be affected by that conduct. However, the more distant the anticipated impact, the more doubtful it may be whether the merged entity will have the incentives to adopt the conduct in question. In such cases, the merged entity may more likely prefer to maximize short-run profits rather than pursue a policy intended to obtain possible, but uncertain, long-run gains.

Article 82

The CFI has provided some limited guidance on the application of the requirement established in the *Tetra Laval* judgments that the deterrent effect of Article 82 must be taken into account in determining the likelihood that the merged firm will engage in anticompetitive bundling or leveraging strategies.

The CFI confirms that the Commission must take into account the potentially unlawful, and thus sanctionable, nature of certain conduct as a factor that might diminish, or even eliminate, incentives for the merged firm to engage in particular conduct. The Commission is not, however, required to establish that the conduct foreseen in the future will actually constitute an infringement of Article 82 or that such an infringement would be detected and punished. The Commission is entitled to limit itself in this regard to a "summary analysis" based on the evidence available to it.

This does not, however, appear to lower the *Tetra Laval* standard of proof, as the CFI further confirmed that the Commission is required to adduce "convincing evidence" in support of any conglomerate effects theories. This might consist, for example, of actual evidence of the parties' intent to engage in the relevant conduct (e.g., based on internal documents of the parties) or economic analysis demonstrating the parties' commercial incentive to do so.

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

An Introduction to Bork (1966)

Douglas H. Ginsburg

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

An Introduction to Bork (1966)

Douglas H. Ginsburg

The Sherman Antitrust Act of 1890, the cornerstone of the U.S. antitrust regime, broadly prohibits contacts, combinations, and conspiracies in "restraint of trade" and makes it unlawful "to monopolize" any line of commerce.[1] The open-textured nature of the Act—not unlike a general principle of common law—vests the judiciary with considerable responsibility for interpretation, the discharge of which requires it to choose among competing values. In this important article,[2] then-Professor Robert H. Bork examined the legislative history of the Sherman Act in search of the U.S. Congress's intent in passing it and, therefore, the policies the judiciary should follow when deciding cases under the Act. Bork was candid about the "difficulties inherent in the very concept" of legislative intent and cautioned against viewing his work "as an attempt to describe the actual state of mind of each of the congressmen who voted for the Sherman Act."[3] Nevertheless, he thought the undertaking justified by the need to counter the judiciary's repeated invocation of values that were unrelated to the debate that had informed congressional enactment of the Sherman

1 *See* 15 U.S.C. §§ 1 and 2.

2 The article was originally published in the *Journal of Law & Economics. See* R.H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L. & Econ. 7 (1966) reprinted in this issue as 2(1) Competition Pol'y Int'l 233-278 (2006). Hereinafter, where the Bork article is cited, the first set of page citations refer to pages in Bork's original article and the second set in parentheses refer to pages in the reprint.

3 Bork, *supra* note 2, at 7 n.1 (at 233 n.1). Bork's caveat is an important one. After all, "[i]t is the *law* that governs, not the intent of the lawgiver." A. Scalia, A Matter of Interpretation 17 (1997).

The author is Chief Judge, U.S. Court of Appeals for the District of Columbia Circuit.

Act and, lacking any legitimate economic rationale, were likely to produce real economic harm.

For example, the Supreme Court of the United States in *Fashion Originators' Guild of America v. FTC* counted protection of "the freedom of action of [Guild] members [not] to reveal to the Guild the intimate details of their individual affairs" among the policies underlying the Sherman Act.[4] Indeed, no lesser light than Judge Learned Hand had asserted that the Congress intended the Sherman Act to achieve certain socio-political aims, such as minimizing the "helplessness of the individual"[5] and ensuring the "organization of industry in small units."[6] Obviously such policies are highly malleable. They can be invoked (or not) to justify almost any result in any situation. Indeed, as Bork pointed out, Judge Hand went so far as to state that in enacting the Sherman Act, the Congress had "delegated to the courts the duty of fixing the standard in each case."[7]

Bork's examination of the text and structure of the Sherman Act against the background of preliminary proposals and draft legislation, statements by senators and representatives, and contemporaneous understandings of constitutional and common law led him to conclude: "The legislative history . . . contains no colorable support for application by courts of any value premise or policy other than the maximization of consumer welfare."[8] By "consumer welfare" Bork meant "the maximization of wealth or consumer want satisfaction,"[9] known today as allocative efficiency, a concept he thought the framers of the Sherman Act clearly grasped even though they did not "speak of consumer welfare with the precision of a modern economist."[10] Bork also explained that maximization of consumer welfare is the common denominator underlying the central prohibitions of the Act, that is, the condemnation of cartel agreements, monopolistic mergers, and predatory business tactics.[11] He explained that legislators used the term "monopolize" to refer only to those three prohibited activities, as opposed to a "monop-

4 312 U.S. 457, 465 (1941).

5 United States v. Aluminum Co. of America, 148 F.2d 416, 428 (2d Cir. 1945).

6 *Id.* at 429.

7 United States v. Associated Press, 52 F. Supp. 362, 370 (S.D.N.Y. 1943), *aff'd*, 326 U.S. 1 (1945).

8 Bork, *supra* note 2, at 10 (at 236).

9 *Id.* at 1 (at 234).

10 *Id.* at 10 (at 236).

11 *Id.* at 11-12, 21-26 (at 237-39, 248-53).

oly," which might arise from superior efficiency.[12] According to Bork, "Only a consumer-welfare value which, in cases of conflict, sweeps all other values before it can account for Congress' willingness to permit efficiency-based monopoly."[13]

When Bork's article was first published in 1966, his thesis was novel. By 1977, it had become the conventional wisdom of the federal courts. That year the U.S. Supreme Court in *Continental T.V., Inc. v. GTE Sylvania Inc.*[14] repudiated the position it had taken only ten years before in *United States v. Arnold, Schwinn & Co*. In the earlier case, the Court had held that a non-price vertical restraint imposed by a manufacturer on a distributor after "title, dominion, or risk" had passed was a per se violation of the Sherman Act,[15] that is, regardless of its actual—and possibly efficient—economic effect.

In *GTE Sylvania Inc.*,[16] a retailer of televisions claimed a manufacturer's limitation on the locations at which the retailer could sell its televisions was a per se violation of the Sherman Act. The U.S. Court of Appeals for the Ninth Circuit, after rehearing the case en banc, had recognized that—as the Supreme Court later put it—the condemnation of "*Schwinn* [was] clearly broad enough to apply" to the facts of the case.[17] Nonetheless, the Ninth Circuit concluded *Schwinn* was not controlling, applied the rule of reason, and endorsed the manufacturer's position that such arrangements "may in some instances *promote*, rather than impede, competition" and, in turn, allocative efficiency.[18] More to the present point, the Ninth Circuit expressly adopted Bork's thesis and rejected the multiplicity of "values" that the Supreme Court for decades had been reading into the Sherman Act:

> "Since the legislative intent underlying the Sherman Act had as its goal the promotion of consumer welfare, we decline blindly to condemn a business practice as illegal per se because it imposes a partial, though perhaps reason-

12 *Id.* at 12, 26-31 (at 238-39, 254-59).

13 *Id.* at 12 (at 238).

14 433 U.S. 36 (1977).

15 388 U.S. 365, 379 (1967).

16 433 U.S. 36 (1977).

17 *Id.* at 46.

18 GTE Sylvania Inc. v. Continental T.V., Inc., 537 F.2d 980, 1000 (9th Cir. 1976) (en banc).

able, limitation on intrabrand competition, when there is a significant possibility that its overall effect is to promote competition between brands."[19]

Two dissenters remained of the view that the legislative history of the Sherman Act "reflect[s] a concern not only with the consumer interest in price, quality, and quantity of goods and services, but also with society's interest in the protection of the independent businessman, for reasons of social and political as well as economic policy."[20]

The Supreme Court affirmed the Ninth Circuit, holding that "[*p*]*er se* rules of illegality are appropriate only when they relate to conduct that is manifestly anticompetitive,"[21] and stating, "[v]ertical restrictions promote interbrand competition by allowing the manufacturer to achieve certain efficiencies in the distribution of his products."[22] In emphasizing allocative efficiency over other values, the Supreme Court implicitly endorsed Bork's thesis. Indeed, in his concurring opinion, Justice White attributed to the Court the view that the Sherman Act is "directed solely to economic efficiency," citing Bork's article as the source of that position.[23]

The significance of the Court's new, Borkian position should not be underestimated. As Professor Timothy Muris has said, "the opinion was a ringing endorsement of the economic approach to antitrust law."[24]

Two years later, in *Reiter v. Sonotone Corp.*,[25] the Supreme Court considered a class action brought under the Clayton Antitrust Act of 1914 by plaintiffs who had purchased hearing aids from a manufacturer they alleged had fixed prices with its rivals and its retailers. Relying this time expressly on Bork's appraisal of the legislative history of the Sherman Act as the "predecessor" of the Clayton Act, the Court concluded the latter Act, in providing a remedy to anyone

19 *Id.* at 1003 (footnote omitted). *See id.* at n.39 ("A study of the legislative history of the Sherman Act 'establish[es] conclusively that the legislative intent underlying the Sherman Act was that courts should be guided exclusively by consumer welfare and the economic criteria which that value premise implies'") (quoting Bork, *supra* note 2, at 11 (at 237)).

20 GTE Sylvania Inc., 537 F.2d at 1019 (Browning, J., joined by Wright, J., dissenting).

21 Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 50 (1977).

22 *Id.* at 54.

23 *Id.* at 69 (White, J., concurring) (citing Bork, *supra* note 2, at 7 (at 233)).

24 T.J. Muris, GTE Sylvania *and the Empirical Foundations of Antitrust*, 68 ANTITRUST L.J. 899, 900 (2001).

25 442 U.S. 329 (1979).

injured in his "business or property," covered "pecuniary injuries suffered by those who purchase goods and services at retail for personal use."[26] Quoting Bork's 1978 book, *The Antitrust Paradox*, in which a version of his article appears as a chapter, the Court declared that the legislative history "suggest[s] that Congress designed the Sherman Act as a 'consumer welfare prescription.'"[27]

In *NCAA v. Board of Regents of University of Okalahoma*,[28] the Court had further occasion to embrace the consumer welfare thesis when it determined the National Collegiate Athletic Association's limitation on the number of televised intercollegiate football games and its fixed-price, exclusive agreements with certain broadcasters violated the Sherman Act. Although the Court noted the arrangement adversely affected competitors' "freedom to compete," it ultimately based its decision squarely on allocative efficiency:

> "Price is higher and output lower than they would otherwise be, and both are unresponsive to consumer preference. This . . . point is perhaps the most significant, since Congress designed the Sherman Act as a consumer welfare prescription. A restraint that has the effect of reducing the importance of consumer preference in setting price and output is not consistent with this fundamental goal of antitrust law."[29]

Thus, by the mid-1980s, Bork's thesis had undeniably changed the Supreme Court's most fundamental understanding of the Sherman Act.

Academics began seriously to challenge Bork only after the Supreme Court had adopted his reading of the legislative history in *Reiter*. From the more than a dozen articles critical of the consumer welfare thesis, there emerge two distinct alternative theories of congressional intent. One, advanced by Professor Robert H. Lande, is that the Congress's chief objective in the Sherman Act was the prevention of "wealth transfers" from consumers to business trusts, forerunners of the large corporations of today.[30] Though he agrees with Bork that some legisla-

26 *Id.* at 343.

27 *Id.* (quoting R.H. Bork, Antitrust Paradox 66 (1978)).

28 468 U.S. 85 (1984).

29 *Id.* at 107-08 (footnotes, citations, and quotation marks omitted).

30 *See* R.H. Lande, *Wealth Transfers as the Original and Primary Concern of Antitrust: The Economic Efficiency Interpretation Challenged*, 34 Hastings L.J. 65 (1982).

tors were concerned with allocative efficiency, Lande maintains that a number of them believed large trusts were generally more efficient than small- and medium-sized businesses. Because the Sherman Act is an "anti-trust" measure, Lande concludes allocative efficiency could not have been the sole value underlying the statute. Instead, he argues the Act was intended to curb the market power of large producers in order to prevent their "extract[ing] wealth from consumers."[31]

Professor Herbert Hovenkamp, although he believes Lande's case is stronger than Bork's,[32] contends the primary purpose of the Sherman Act was the protection of small business, not of consumers. The legislative history of the Sherman Act and attendant political circumstances, he believes, "suggest that the interest groups that communicated their concerns to Congress most effectively were small producers."[33] Hovenkamp concludes that the Congress acted neither solely on the basis of efficiency nor only in order to benefit consumers, but rather primarily to avert "various kinds of injury to competitors . . . flow[ing] mainly from the lower costs of more efficient rivals."[34]

The challenges to Bork's thesis lodged by Lande and Hovenkamp are representative of the academy as a whole. One commentator goes so far as to claim that Bork's interpretation "has been almost universally rejected by antitrust scholars."[35] Yet the academy has failed to persuade the judiciary, and Bork's consumer welfare thesis has become one of his many enduring contributions to U.S. antitrust law.

Regardless whether Bork's assessment of the legislative history of the Sherman Act is correct, the Supreme Court's endorsement of allocative efficiency as the fundamental value underlying the antitrust laws has had important consequences. First, as a matter of administrability, the consumer welfare thesis has substantially ameliorated the practical problem of having courts choose among multiple, incommensurable, and often conflicting values. Even one of Bork's sharpest critics agrees. Professor Christopher Grandy, who concludes "the legislative history of the Sherman Act fails to support the consumer-welfare hypothesis," nevertheless acknowledges that Bork's thesis "provides a clear and cogent set

31 *Id.* at 93.

32 *See* H. Hovenkamp, *Antitrust's Protected Classes*, 88 MICH. L. REV. 1, 24 (1989).

33 *Id*.

34 P. AREEDA & H. HOVENKAMP, ANTITRUST LAW ¶ 101, at 11 (2d ed. 2000).

35 P.J. Hammer, *Antitrust Beyond Competition: Market Failures, Total Welfare, and the Challenge of Intramarket Second-Best Tradeoffs*, 98 MICH. L. REV. 849, 905 n.150 (2000).

of rules that courts can apply in antitrust cases, and no other view of antitrust accomplishes that task as well."[36]

Second, judicial adoption of Bork's thesis has nearly put an end to the efforts of counsel and the propensity of lower courts to manipulate outcomes by invoking highly plastic, subjective values of the sort instanced by Judge Hand. Third, by applying a single standard rooted in economic analysis, court decisions have become less arbitrary and more predictable. No longer must businesses make decisions without knowing the standard by which their actions, if challenged by the courts, will later be judged. Finally, judicial endorsement of the consumer welfare thesis has no doubt lead to a more efficient allocation of scarce resources, thereby increasing the wealth of the nation. Had Bork not written the following article, these salutary developments might still be in the offing.

36 C. Grandy, *Original Intent and the Sherman Antitrust Act: A Re-Examination of the Consumer-Welfare Hypothesis*, 53 J. Econ. Hist. 359, 373 (1993).

VOLUME 2 | NUMBER 1 | SPRING 2006

Competition Policy International

Legislative Intent and the Policy of the Sherman Act

Robert H. Bork

Originally published in *Journal of Law and Economics*.

Published in *Competition Policy International* (print ISSN 1554-0189, online ISSN 1554-6853), Spring 2006, Vol. 2, No. 1.
Competition Policy International is a free publication. To order or download additional copies, visit eSapience.org.

Legislative Intent and the Policy of the Sherman Act

Robert H. Bork

Originally published in *Journal of Law and Economics,* Vol. 9, No. 7, 1966.

Despite the obvious importance of the question to a statute as vaguely phrased as the Sherman Act, the federal courts in all the years since 1890 have never arrived at a definitive statement of the values or policies which control the law's application and evolution. The question of values, therefore, remains central to controversy about this basic law and its interpretation. More than one factor bears upon the answer to the question. Courts do not and should not, for example, attempt to administer any policy a legislature may seek to thrust upon them.[1] Nevertheless, a starting point is the question of legislative intent.[2] In this paper I propose to examine that question. My conclusion, drawn from the evidence in the Congressional Record, is that Congress intended the courts to implement (that is, to take into account in the decision of cases) only that value we would today call consumer welfare. To put it another way, the pol-

1 I have discussed elsewhere the problem of judicially administrable standards under the Sherman Act. See Bork, The Rule of Reason and the Per Se Concept: Price Fixing and Market Division, 74 Yale L.J. 775, 829-847 (1965). My conclusion there is that consumer welfare provides a proper standard and that most other suggested policies do not. In that article, however, I seriously underestimated the clarity of the legislative intent behind the Sherman Act which a closer study of the full record reveals.

2 The attribution of any intent to a legislature involves a number of problems and assumptions. My justification for ignoring the difficulties inherent in the very concept of legislative intention lies primarily in the fact that courts and lawyers do regularly "find," describe, and rely upon such intentions. What I have to say in this paper, therefore, should not be taken as an attempt to describe the actual state of mind of each of the congressmen who voted for the Sherman Act but merely as an attempt to construct the thing we call "legislative intent" using conventional methods of collecting and reconciling the evidence provided by the Congressional Record.

This paper relies upon research undertaken for the American Enterprise Institute for Public Policy Research which is designed to eventuate in a study of the legislative history of the major antitrust statutes. I wish to acknowledge my gratitude to Professor Ward S. Bowman of the Yale Law School for his comments on the penultimate draft of this article.

icy the courts were intended to apply is the maximization of wealth or consumer want satisfaction. This requires courts to distinguish between agreements or activities that increase wealth through efficiency and those that decrease it through restriction of output.

Failure to settle the issue of values has led inevitably to a degree of irresponsibility in the judicial process. Often a court will apply a value in deciding a Sherman Act case without explaining either the selection of the value or the method of its application to the facts. A value will be announced as pertinent with a confidence that is matched only by the mystery that shrouds its derivation. A very specific decision is then whelped from the value premise without benefit of midwifery by any visible minor premise. One is tempted, and perhaps occasionally entitled, to suspect that such a suddenly appearing value is a *dues ex machina* by which the court rescues itself from the perplexing tasks of economic analysis and judgment that rigorous adherence to a consumer-welfare value premise would sometimes require.

It would be possible to illustrate the use of values other than consumer welfare in a number of cases, but the fact of judicial reliance upon such values is surely not in dispute,[3] and excerpts from two well-known opinions of Judge Learned Hand may therefore suffice to illustrate the point. Values other than consumer welfare apparently played large roles in Judge Hand's reasoning in both the *Alcoa* and *Associated Press* cases.

In *Alcoa*, the Court of Appeals for the Second Circuit judged illegal Aluminum Company of America's large market position in virgin aluminum ingot. In an assertion seemingly important to his argument, Judge Hand said:

> "We have been speaking only of the economic reasons which forbid monopoly; but . . . there are others, based upon the belief that great industrial consolidations are inherently undesirable, *regardless of their economic results*. In the debates in Congress Senator Sherman himself . . . showed that among the purposes of Congress in 1890 was a desire to put an end to great aggregations of capital because of the helplessness of the individual before them."[4] (Emphasis added.)

3 Among many examples that might be cited of opinions employing values other than consumer welfare are Chicago Board of Trade v. United States, 246 U.S. 231 (1918) (Brandeis, J.); Associated Press v. United States, 326 U.S. 1, 25-29 (1945) (concurring opinion of Frankfurter, J.); United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948) (Douglas, J.); Fashion Originators' Guild v. Federal Trade Commission, 312 U.S. 457 (1941) (Black, J.).

4 United States v. Aluminum Co. of America, 148 F.2d 416, 428 (2d Cir. 1945). In a footnote Judge Hand supports his assertion by two quotations from Senator Sherman and a page citation to Senator George. These passages are analyzed to determine whether they support Hand's thesis at pp. 39-42, *infra*.

Without pausing to explain what the noneconomic helplessness of the individual might consist of, what category of individuals was involved, or how the concept applied to the facts of the case before him, Judge Hand moved on to another formulation of noneconomic values supposedly embedded in the statute:

> "Throughout the history of these statutes [the antitrust laws, including the Sherman Act] it has been constantly assumed that one of their purposes was to perpetuate and preserve, for its own sake and in spite of possible cost, an organization of industry in small units which can effectively compete with each other."[5] (Emphasis added.)

This passage was followed immediately by: "We hold 'Alcoa's' monopoly of ingot was of the kind covered by Sec. 2 [of the Sherman Act]."

The italicized phrases in each of the foregoing quotations indicate that Judge Hand was asserting that the nebulous values he derived from the legislative history, or from prevalent assumptions about the legislative history, were powerful enough to require a court to override considerations of consumer welfare. He did not inform us whether that was true in all cases where the "economic" value of consumer welfare conflicted with these other values or, if not, how to predict the cases in which one or the other of these conflicting values would take precedence.

But Judge Hand went further even than this. In his *Associated Press* opinion he asserted that the Fifty-first Congress had given the federal courts virtual carte blanche to choose the values they would implement through the Sherman Act. Approaching his topic through a rapid survey of antitrust doctrine and using a cluster of trade association cases for his springboard, Judge Hand said:

> "[T]he injury imposed upon the public was found to outweigh the benefit to the combination, and the law forbade it. We can find no more definite guide than that."

5 148 F.2d at 429. Earlier in the opinion Judge Hand said that Congress "did not condone 'good trusts' and condemn 'bad' ones; it forbade all. Moreover, in so doing it was not necessarily actuated by economic motives alone. It is possible, because of its indirect social or moral effect, to prefer a system of small producers, each dependant for his success upon his own skill and character, to one in which the great mass of those engaged must accept the direction of a few. These considerations, which we have suggested only as possible purposes of the Act, we think the decisions prove to have been in fact its purposes." *Id*. at 427. It was not made clear how the factual question of what legislators intended is proven one way or the other by judicial decisions. Judge Hand's speculations concerning possible purposes can only be tested against the legislative record.

> Certainly such a function is ordinarily "legislative"; for in a legislature the conflicting interests find their respective representation, or in any event can make their political power felt, as they cannot upon a court. . . . But it is a mistake to suppose that courts are never called upon to make similar choices: i.e., to appraise and balance the value of opposed interests and to enforce their preference. The law of torts is for the most part the result of exactly that process, and the law of torts has been judge-made, especially in this very branch. Besides, even though we had more scruples than we do, we have here a legislative warrant, because Congress has incorporated into the Anti-Trust Acts the changing standards of the common law, and by so doing has delegated to the courts the duty of fixing the standard for each case.[6]

The liberating potential of this judicial equivalent of free verse or "tennis with the net down" was demonstrated as Judge Hand went on to note that Associated Press' by-laws made attainment of membership more difficult for newspapers in competition with present members, that non-members were disadvantaged by being unable to get Associated Press news, that the First Amendment expresses an important value in our society, and, finally, that this value weighed against the Sherman Act legality of the by-laws.[7] The method by which Judge Hand moved from First Amendment values to the illegality of the by-laws left a great deal—in fact, almost everything—to be desired. Passing that, however, the propriety of Judge Hand's consideration of First Amendment values at all demands that Congress' "incorporation" of "the common law" into the Sherman Act have been intended to delegate a value-choosing role to the federal judiciary.

I do not wish to focus upon Judge Hand. He is cited here merely as an authoritative and persuasive spokesman for positions which are widely held and which I wish to dispute. There would be little point in reviewing here all of the positions that have been advanced concerning the broad social, political, and ethical mandates entrusted to the courts through the Sherman Act, or in naming the persons who have urged them, for there is not a scintilla of support for most such views anywhere in the legislative history. The only value other than consumer welfare which is even suggested by the record is protection of small businessmen, but, as will be argued, that value was given only a complementary and not a conflicting role. The legislative history, in fact, contains no colorable support for application by courts of any value premise or policy other than the maximization of consumer welfare. The legislators did not, of course, speak of consumer welfare with the precision of a modern economist but their meaning was unmistakable.

A point which requires emphasis at the outset is the distinction, alluded to above, between conflicting and complementary values. I recognize that many of the legislators who voted for the Sherman Act may have had values in mind in

6 United States v. Associated Press, 52 F. Supp. 362, 370 (1943).

7 *Id.* at 368-373.

addition to or other than consumer welfare. There was, for example, repeated expression of concern over the injury trusts and railroad cartels inflicted upon farmers and small businessmen. It by no means follows, however, that Congress intended courts to take such concerns into account under the statute. A legislator may be moved to vote for a statute by his perception that it will affect a range of values which are not reflected in the criteria that the law requires the courts to use. In the case of the Sherman Act it seems quite clear that this was the situation. Not only was consumer welfare the predominant goal expressed in Congress but the evidence strongly indicates that, in case of conflict, other values were to give way before it. This means that such other values are superfluous to the decision of cases since none of them would in any way alter the result that would be reached by considering consumer welfare alone. For a judge to give weight to other values, therefore, can never assist in the correct disposition of a case and may lead to error. In short, since the legislative history of the Sherman Act shows consumer welfare to be the decisive value it should be treated by a court as the only value.

. . . THE PURPOSE OF ANTITRUST LEGISLATION WAS CONSUMER WELFARE AND THAT THAT POLICY WAS TO GUIDE THE COURTS.

Following these guidelines, then, the following arguments, which will be supported by evidence from the record, seem to me, when taken together, to establish conclusively that the legislative intent underlying the Sherman Act was that courts should be guided exclusively by consumer welfare and the economic criteria which that value premise implies.

1. Both in the bills introduced and in the debates there are a number of explicit statements that the purpose of antitrust legislation was consumer welfare and that that policy was to guide the courts.

2. The rules of law which Congress foresaw are inconsistent with any value premise other than consumer welfare. Congress contemplated that the statute would strike at three basic phenomena: cartel agreements; monopolistic mergers; and predatory business tactics.

    a. A rule of *per se* illegality for cartel agreements (agreements whose purpose is not to produce efficiency but merely to eliminate competition) discloses a policy judgment that firms should fare well or ill according to the standards consumers impose in a competitive marketplace. Such a rule leaves a court no discretion to weigh other values which might legitimate the cartel: for example, the preservation of existing small businessmen, or the welfare of those businessmen who would prefer a shorter work day if their rivals would agree to close down too. The flat prohibition of cartel agreements which Congress envisaged seems fully consistent only with the idea that output should not be artificially restricted, and that desire is in turn explained only by a concern for consumer well-being.

b. A rule against monopolistic mergers, taken by itself, may appear less unequivocally to imply a consumer welfare rationale. The fact that the rule is phrased in terms of monopoly rather than absolute size suggests such a rationale, but the rule could conceivably reflect values of the sort Judge Hand sketched in his *Alcoa* opinion. The argument for this rule in Congress, however, shows that it derived in large measure from a desire to protect consumers from monopoly extortion. Insofar as other classes, such as small producers who sold to or bought from monopolists, were to be benefited, that benefit was not seen as conflicting with the consumer-welfare rationale but rather as reinforcing it. Where producer and consumer welfare might come into conflict, as will be seen under point 3 below, Congress chose consumer welfare as decisive.

c. A similar policy ambiguity may seem at first glance to accompany a rule outlawing predatory business practices. A law against "unfair" commercial tactics could be rooted in moral or humane considerations, a wish to introduce Marquis of Queensbury rules into the commercial arena, either for the sake of the combatants or of the spectators. An alternative hypothesis is provided, however, by an economic theory widely held then as now. Business firms with large capital or low ethics were thought capable of gaining or preserving monopoly positions by crushing rivals with tactics, such as selling below cost, which do not reflect superior efficiency. This theory leads the legislator who entertains it to outlaw injury to competitors only when it is a step toward monopoly and does not result from the exercise of efficiency. The terms of the arguments made in Congress as well as the attitude of Congress toward efficiency indicate that this second hypothesis explains the congressional antipathy to "unfair" practices. The rule thus rests on a consumer-welfare rationale.

3. Congress was very concerned that the law should not interfere with business efficiency. This concern, which was repeatedly stressed, was so strong that it led Congress to agree that monopoly itself was lawful if it was gained and maintained only by superior efficiency. Thus the desire to protect small firms from annihilation by monopoly-minded rivals did not extend an inch beyond the bounds of the consumer-welfare rationale. Small producers would be equally threatened by a rival on its way to monopoly through superior efficiency. The noneconomic helplessness of the individual to which Judge Hand referred would, moreover, seem to be the same before any monopoly, no matter how gained. Only a consumer-welfare value which, in cases of conflict, sweeps all other values before it can account for Congress' willingness to permit efficiency-based monopoly. To break up such monopolies because rivals could not meet their prices would be to impose lower output and higher prices upon consumers.

4. That Congress did not wish courts to apply criteria expressing values other than consumer welfare is also strongly suggested by its preferred method of dealing with situations in which consumer welfare was not to be controlling. The primary examples were farm and labor organizations. Most of the congressmen who spoke to this issue favored the complete exemption of such organizations from the coverage of the statute. Senator Edmunds, who appears to have played the primary role in drafting the bill which became the Sherman Act, wished to include such groups within the law's sweep. The Act as passed was silent on the issue. It may be uncertain, therefore, whether Congress had an intention on this issue and, if it did, what that intention was. But it is clear that those who did not wish farm and labor organizations judged by consumer-welfare criteria adopted the technique of exempting them from the bill altogether. No one suggested that the matter be handled by letting the courts balance the values that these congressmen thought were in play. This raises a fairly strong inference that no values other than consumer welfare were to be considered in those cases which were intended to come within the statute's coverage.

5. Given the narrow view of the commerce power that prevailed in 1890 it is extremely unlikely that the Fifty-first Congress intended to give the courts the power to make broad social or political decisions through the Sherman Act. The federal commerce power was circumscribed not merely by the wide category of commerce that was intrastate but also by its nature as a commercial power. It was generally assumed, that is, that the ends to be accomplished by the exercise of the commerce power must themselves be of a commercial nature. This assumption would not impose a consumer-want-satisfaction rationale upon the statute—the category of commercial purposes comprises more than that—but it does tend to rule out an intention to achieve the broad noncommercial goals that are sometimes attributed to the Sherman Act. The discussions of the commerce power in Congress, as well as the phrasing given the statute by the Judiciary Committee, bear out this thesis.

6. Congress recognized that broad areas of discretion were being delegated to the courts but not one speaker suggested that that discretion included the power to consider any values other than consumer welfare. Senator Sherman, on the other hand, was as explicit as could be desired that the criteria by which the delegation was to be controlled were those relating to consumer welfare. The statute's incorporation of a highly artificial version of "the common law" further demonstrates the consumer-welfare limits of the discretion delegated to the courts.

7. The complete absence of any expression of values which conflict with consumer welfare among those urging antitrust legislation is itself compelling evidence that no such values were intended. Those few legislators who urged that producer welfare override consumer interests in some cases did do, significantly, in opposing the bills drafted by Senators Sherman and Reagan.

Finally, an objection to the thesis advanced here will be discussed. This consists of the argument that the legislative intent underlying the statute is essentially unknowable because the Judiciary Committee draft which was enacted was totally different from Sherman's and Reagan's drafts which were discussed. It can be shown, however, that the policies of the drafts were the same so that the debates are fully applicable to the Act as it stands today.

The narrative of the drafting, discussion, and enactment of the Sherman Act has been told by others.[8] I will give only the briefest outline here.

Senator Sherman introduced S.1 in December 1889.[9] It was called up for debate before the Senate in Committee of the Whole on February 27, 1890, and subjected to a detailed, scathing attack upon its constitutionality and efficacy. The Finance Committee, of which Sherman was the leading member, responded by reporting a modified version of S.1 on March 18. Neither the criticisms nor the modifications concerned the bill's criteria for illegality. Debate on the modified bill began on Friday, March 21, with a lengthy explanation of S.1 and its policies by Sherman. The process of discussion and amendment continued through Thursday, March 27, when the bill was referred to the Judiciary Committee for redrafting. The Judiciary Committee's redraft of S.1, which ultimately became the Sherman Act, was reported back on April 2 and passed the Senate, by a vote of 52 to 1, on April 8. House debate followed and a proposed House amendment, with a Senate amendment in response, led to two conferences before both houses receded and the bill was enacted as it had first come from the Senate. President Harrison signed the bill on July 2, 1890.

Prior to the redraft of S.1 by the Senate Judiciary Committee the Senate in Committee of the Whole had adopted so many amendments in the nature of additions that the bill had become a monstrosity. The more important additions for our purposes were those proposed by Senator Reagan (D., Texas), which dealt with the same problems as Sherman's bill, and Senator Ingalls (R., Kansas),

8 See, particularly, Thorelli, The Federal Antitrust Policy 164-214 (1954).

9 Sherman had introduced in the preceding Congress a resolution directing the Committee on Finance to inquire into and report on measures to control anticompetitive agreements and combinations. Apparently in order to justify the delegation of such a question to the Committee on Finance, of which Sherman was the most influential member, rather than to the more appropriate Committee on the Judiciary, to which he did not belong, the resolution directed that proposed measures be taken up in connection with any bill raising or reducing revenue. The Senate adopted the resolution without debate. Later in that same Congress Sherman introduced a bill, S.3445, which was referred to the Finance Committee and reported back in amended form. The Senate discussed this bill but took no action on the subject during the 50th Congress. This bill was reintroduced by Sherman as S.1 in the 51st Congress. The provisions of the resolution and of Sherman's bill are discussed *infra* at note 11.

which placed a prohibitive tax upon dealings in options and futures. These and a host of minor amendments made the bill so complex as to be incomprehensible. It was for this reason, as well as because of widespread doubt concerning the constitutionality of the various measures as framed, that the Judiciary Committee was asked to write a new draft.

With this outline of the order of events in mind we may proceed to consider the evidence of Congress' intent.

I. Explicit Policy Statements

The views of Senator Sherman (R., Ohio), are crucial to an understanding of the intent underlying the law that bears his name. Sherman was the prime mover in getting antitrust legislation considered and pressed through the Senate. He was also by far the most articulate spokesman for antitrust in Congress. It will be seen, moreover, that though Sherman's bill was completely rephrased by the Judiciary Committee, of which he was not a member, the final bill, in its substantive policy aspects, embodied Sherman's views.

Sherman's views on the policy to be served by antitrust legislation are clear. They appear on the face of the bill he drafted and reported from the Committee on Finance, S.1. Section 1 of that bill declared illegal two classes of "arrangements, contracts, agreements, trusts, or combinations": (1) those "made with a view, or which tend, to *prevent full and free competition*," and (2) those "designed, or which tend, *to advance the cost to consumer*" of articles of commerce.[10] Sherman employed these two criteria of illegality in every measure he presented to the

10 The complete section read:

> That all arrangements, contracts, agreements, trusts, or combinations between persons or corporations made with a view or which tend to prevent full and free competition in the importation, transportation, or sale of articles imported into the United States, or in the production, manufacture, or sale of articles of domestic growth or production, or domestic raw material that competes with any similar article upon which a duty is levied by the United States, *or* which shall be transported from one State or Territory to another, and all arrangements, contracts, agreements, trusts, or combinations between persons or corporations designed or which tend to advance the cost to the consumer of any such articles are hereby declared to by against public policy, unlawful, and void.

The second section of the bill provided for private suits to recover the sum paid for any goods "included in or advanced in price by said combination." The third made participation in a prohibited arrangement, etc., a criminal offense and punishable by a fine of not more than $10,000, imprisonment for not more than five years, or both. 21 Cong. Rec. 1765 (1890). Sherman dropped the third section when he reported the Finance Committee's modified version on March 21, 1890. *Id.* at 2455.

Senate.[11] The first test, which subjects all firms to market forces, is hardly a means of preserving social values that consumers are not willing to pay for. It can be reconciled only with a consumer-welfare policy. The second test is even more explicit. The touchstone of illegality is raising prices to consumers. There were no exceptions. Sherman wanted the courts not merely to be influenced by the consumer interest but to be controlled completely by it.

SHERMAN WANTED THE COURTS NOT MERELY TO BE INFLUENCED BY THE CONSUMER INTEREST BUT TO BE CONTROLLED COMPLETELY BY IT.

Sherman's speeches in support of his fill fully bear out this reading. He said, for example, that his bill sought "only to prevent and control combinations made with a view to prevent competition, or for the restraint of trade, or to increase the profits of the producer at the cost of the consumer;"[12] that a combination which embraced "the great body of all the corporations engaged in a particular industry" tended "to advance the price to the consumer," and was "a substantial monopoly injurious to the public;"[13] and, speaking

11 These tests were used in S.3445, Sherman's bill in the 50th Congress, and in every draft of S.1 which he offered to the 51st. The same policy orientation is shown by the terms of the resolution he offered and the Senate adopted in the 50th Congress:

> *Resolved*, That the Committee on Finance be directed to inquire into and report, in connection with any bill raising or reducing revenue that may be referred to it, such measures as it may deem expedient to set aside, control, restrain, or prohibit all arrangements, contracts, agreements, trusts, or combinations between persons or corporations, *made with a view, or which tend to prevent free and full competition* in the production, manufacture, or sale of articles imported into the United States, or which, against public policy, *are designed or tend to foster monopoly or to artificially advance the cost to the consumer* of necessary articles of human life, with such penalties and provisions, and as to corporations, with such forfeitures, as will tend *to preserve freedom of trade and production, the natural competition of increasing production, the lowering of prices by such competition*, and the full benefit designed by and hitherto conferred by the policy of the government to protect and encourage American industries be levying duties on imported goods.

(Emphasis added.) 19 Cong. Rec. 6041 (1888).

The first two italicized passages above indicate, as the text discusses, the evils which Sherman wished to avert. The third italicized passage shows the benefits he wished to secure. The evils are described as prevention of competition, monopoly, and the artificial advancement of prices to consumers. The benefits are freedom of trade and production, increasing production, and the lowering of prices by the competition of increasing production. It could hardly be clearer that Sherman wanted to stop restrictions of output and permit efficiency. These are goals, as the text will argue, which can only be related to consumer welfare. (The last sentence of this resolution reflects the Republican's contention that protective tariffs were beneficial to consumers as well as producers. The inconsistency of this argument with the arguments for antitrust was either not apparent to Sherman and the Republican majority-though pointed out incessantly by the Democrats-or did not perturb them. In any event, the tariff approach to domestic competition was never suggested by Sherman or others who supported his antitrust objectives.)

12 21 Cong. Rec. 2457 (1890).

13 *Ibid.*

of the trusts, "If they conducted their business lawfully, without any attempt by these combinations to raise the price of an article consumed by the people of the United States, I would say let them pursue that business."[14]

Though an economist of our day would describe the problem of concern to Sherman differently, as a misallocation of resources brought about by a restriction of output rather than one of high prices, there is no doubt that Sherman and he would be talking about the same thing. Indeed, Sherman demonstrated more than once that he understood that higher prices were brought about by a restriction of output. In defending his bill's constitutionality, for example, he asked, wholly rhetorically, whether Congress had not the power to "protect commerce, nullify contracts that restrain commerce, turn it from its natural courses, increase the price of articles, and thereby diminish the amount of commerce?"[15] This and other remarks suggest that Sherman and his colleagues identified the phrase "restraint of commerce" or "restraint of trade" with "restriction of output."[16] If this identity can be carried over to the wording of the Sherman Act, as I believe it can, the meaning of that statute becomes clear and its consumer orientation indisputable.

After Sherman in importance in the legislative career of the statute stand the members of the Senate Judiciary Committee which reworded the bill after most of the debate had taken place. The members of that Committee were Edmunds (R., Vermont); Hoar (R., Massachusetts); Ingalls (R., Kansas); Evarts (R., New York); Wilson (R., Iowa); Coke (D., Texas); Vest (D., Missouri); George (D., Mississippi); and Pugh (D., Alabama). Of these men, four—George, Coke, Vest, and Pugh who comprised the Democratic minority—gave explicit evidence that they agreed with the consumer-welfare rationale offered by Sherman. Of the five Republicans, none gave evidence of disagreement with that policy and several gave indirect evidence, to be discussed in later sections of this paper, that they agreed.

George was a vociferous critic of the constitutionality and efficacy of Sherman's bill on such issues as the inability of the commerce to deal with manufacturing and the difficulties of proving intent,[17] but his agreement with that bill's value premise is shown by the bill he drafted. George's bill employed the

14 *Id.* at 2569.

15 21 Cong, Rec. 2462 (1890).

16 S.1 was entitled a bill "to declare unlawful trusts and combinations in restraint of trade and production." The phrase "in restraint of production" seems hardly to bear any other reading than "in restriction of production," and this throws light upon the companies phrase "in restraint of trade." Apparently Sherman thought of production and trade as separable phases of the economic process, and the two phrases together are subsumed within the modern phrase "restriction of output." The idea that restriction of output was at the root of the problem to be dealt with, was expressed by others as well. Senator Pugh and Representative Heard both expressed that idea. See pp. 18-20 *infra*.

17 See, for example, 21 Cong. Rec. 1765-1772 (1890).

same tests for illegality as Sherman's—the prevention of competition and the advancement of costs to consumers.[18] George's speeches showed him to be concerned with the effect of the trusts upon the small producers who sold to or bought from them, but his bill confirms the internal evidence in his speeches that he did not wish the courts to protect small producers at the expense of consumers. George's concern for producers was entirely complementary to his concern for consumers.[19]

Coke offered his own bill, very similar to Reagan's, and Reagan's bill, as will be shown,[20] appeared to reflect the same policies as Sherman's bill. But Coke criticized Sherman's draft for omitting criminal sanctions. His argument that private damage suits would not provide adequate relief confirms his agreement with Sherman concerning the policy the law should serve:

> "How would a citizen who has been plundered in his family consumption of sugar by the sugar trust, or in his consumption of cotton-bagging under the trust covering that indispensable article, or in is consumption of iron or steel by the iron and steel trust recover his damages under the clause? It is simply an impossible remedy offered to him. . . . If the party damnified . . . were a great corporation, a wealthy association, it could employ lawyers and perhaps be able to show some direct damage, but how could the consumers of the articles produced by these trusts, the vast mass of our people—the individuals—go about showing the damages they had suffered? . . . I think the constituents of all of us, the consumers of products which are raised and manufactured in the country, would be absolutely without a remedy under the bill of the Senator from Ohio."[21]

Vest accepted Sherman's goals but doubted the effectiveness of his bill, saying, after he had heard Sherman's lengthy exposition of his views, that if the bill "would effect what he [Sherman] claims for it, I should vote and speak for it until my strength was exhausted in this Chamber."[22] He preferred Coke's bill as likely

18 *Id.* at 96 and 2657.

19 George's speeches are analyzed at pp. 40-42 *infra*.

20 See pp. 21-22 and notes 57, 81 and 82 *infra*.

21 21 Cong. Rec. 2615 (1890).

22 21 Cong. Rec. 2570 (1890). Earlier Vest had said, "I sympathize with the objects of the Senator from Ohio. . . . [B]ut in my judgment to pass a law which the Supreme Court would declare to be unconstitutional is simply to invite additional disaster." *Id.* at 2467.

to prove more effective.[23] Vest's acceptance of the consumer-welfare rationale was also shown by his argument that the real remedy for the evil of the trusts was the elimination of the protective tariff because "We know very well that competition always reduces prices." He said it was no argument for tariffs, even if it were true, that steel rails were as cheap in England as in the Unites States: "I say if you let these two manufacturing interests compete together and create competition, you then secure lower prices to the consumer."[24] He spoke of American manufacturers coming together to "create these combines at the expense of the consumer in order to enhance their own profits."[25]

Pugh supported Sherman's bill on a consumer-welfare rationale and perceived the connection between artificially raised prices and restriction of output:

> "[T]he existence of trusts and combinations to limit the production of articles of consumption entering into interstate and foreign commerce for the purpose of destroying competition in production and thereby increasing prices to consumers has become a matter of public history, and the magnitude and oppressive and merciless character of the evils resulting directly to consumers and to our interstate and foreign commerce from such organizations are known and admitted everywhere. . . ."[26]

Two other senators not on the Judiciary Committee—Gray (D., Delaware) and Teller (R., Colorado)—also stated explicitly that antitrust legislation should serve consumer welfare. Gray did so by introducing an amendment which employed the same consumer-interest tests for illegality as Sherman's bill.[27] Teller disclosed his policy objectives when he stated he might vote for Sherman's bill, though he was "not very much moved by it" because of its lack of an adequate remedy:

23 *Id.* at 2570-2571.

24 *Id*. at 2466.

25 *Ibid.*

26 *Id.* at 2558. Pugh quoted the first section of Sherman's bill-which dealt with agreements and combinations preventing full and free competition or advancing prices to consumers-and asserted that such arrangements violated the public policy of the United States. *Ibid.*

27 *Id.* at 2657. Gray offered as an amendment the bill originally drafted by George and introduced by him as S.6. This bill employed Sherman's criteria for illegality but substituted as remedies a disability to sue for certain rights in the federal courts and a power and duty in the President to suspend all customs duties and import taxes on articles of the type involved in the described agreement or combination. Gray preferred this bill because he thought Sherman's unconstitutional. *Ibid.*

> "Now, how does this bill reach the great evil against which it is aimed? The Standard Oil Trust has been spoken of . . . . But what can we do about it? We do not dissolve the corporation. What do we do? Anybody who is damaged can sue them. When they interfere with somebody who has sunk a well in Ohio and they run down the price of oil until they shut him up, he may have his remedy against them. But that is not what we are complaining of. We are complaining that Standard Oil Company has a tendency to reduce and destroy competition, and thereby, by destroying competition to put up improperly the price of oil. Who suffers by that? The sixty-five millions of people in the United States who use oil; and how do they suffer? How much damage have they sustained? It is inconsequential individually, but great to the whole mass of the people."[28]

In this passage Teller also shows that predatory attacks by the trusts upon their smaller rivals were not to be outlawed simply to preserve competitors but because of the effect of the resulting monopoly upon consumers.

The debate in the House of Representatives contains similar evidence of the purpose of the Sherman Act, though the debate there was shorter and less enlightening concerning the question of values than was the debate in the Senate. One of the clearest statements of the evil which the bill was designed to cure was made by Representative Heard (D., Missouri) in his excoriation of the "dressed-beef combine":

> "[T]his giant robber combination, while perhaps the most damaging of all of its class to the interests of our people, is only one of many which by their methods extort millions from the citizens of this Republic without adding one cent of value to our productions or one iota of increase to our prosperity. In fact, the very object of these giant schemes of combined capital is not to increase the volume of supply, and thus lessen the cost of any useful commodity, but rather to repress, reduce, and control the volume of every article that they touch, so that the cost to consumers is increased while the expenditure for production if lessened, and thereby their profit secured."[29]

28 *Id.* at 2571.

29 21 Cong. Rec. 4101 (1890). He continued: "We know that by such means the trusts which control the markets on sugar, nails, oils, lead, and almost every other article of use in the commerce of this country have advanced the cost of such articles to every consumer, and that without rendering the slightest equivalent therefor these illegal conspiracies against honest trade have stolen untold millions from the people." *Ibid.*

Heard clearly envisaged the law as one which would prohibit market control that led to restriction of output. He correctly stated that such restriction injured those who sell to trusts as well as consumers,[30] but his concern for such small producers was limited to the restriction-of-output situation and thus did not contradict or add to his consumer-welfare rationale. Statements by Representatives Culberson[31] (D., Texas); Wilson[32] (D., West Virginia); Anderson[33] (R., Kansas); Fithian[34] (D., Illinois); and Taylor[35] (R. Ohio), also indicate that they viewed consumer welfare as the policy of the legislation. Culberson reported the Senate bill favorably from the House Judiciary Committee and Taylor was the chairman of that committee.

30 *Ibid.*

31 *Id.* at 4089, 4090.

32 Wilson devoted his speech primarily to the encouragement given trusts by the protective tariff, but he had printed in the Congressional Record some of his newspaper articles which show that his objection to trusts was their tendency to raise prices by limiting supplies of the articles which they controlled. *Id*. at 4096-4097.

33 Anderson opposed railroad rate pools because of their adverse effect upon farmers and consumers. He rejected the argument that pools should be legal if the rates agreed upon were "just and reasonable":

> The question is whether the people shall be protected by the safeguard of competition between carriers, as they are by competition between merchants, or whether we shall legalize combinations so that the railroads may hereafter charge whatever they see fit in defiance of common law and justice.
>
> The gentleman from Vermont talks about "indiscriminate competition" between railroads. What about "indiscriminate competition" between merchants, or between lawyers, or between doctors, or between mechanics?
>
> Does anybody say you should pass a law preventing "indiscriminate competition" between merchants? Not at all. But when these high and holy railroad millionaires come here . . . , then for some mysterious reason we are called upon to legalize their pools, to "regulate" competition between them lest they hurt each other . . . .

Cong. Rec. 5959 (1890).

34 Fithian, in a speech apparently greatly expanded in the printed version, spoke of the need for relief for his farmer constituents but he was also concerned with the impact of trusts upon consumers and phrased his arguments in those terms: "Competition when left free, and when combinations are not formed to prevent the operation of natural laws, will regulate the price of every commodity and will bring the price down to the level of an honest profit."21 Cong. Rec. 4102 (1890).

35 Taylor, the Chairman of the House Judiciary Committee which recommended passage of S.1 as it came from the Senate, opposed trusts because of their injurious effect upon both farmers and consumers. Of the beef trust he said: "This monster robs the farmer on the one hand and the consumer on the other. This bill proposes to destroy such monopolies, such destructive tyrants . . . ." *Id.* at 4098. Even when defending the protective tariff, a topic that occupied many of the speakers, Taylor did so on the ground that it created lower prices beneficial to consumers. Despite the fallacy of his argument, this demonstrates that Taylor, like most other legislators, was not willing to argue for a policy of preferring producers to consumers.

Additional evidence of the intent of the House of Representatives is provided by the bills introduced there. In the Fifty-first Congress no antitrust bill introduced in the House—or the Senate, either, for that matter—mentioned a value other than consumer welfare. I have counted ten House bills which related explicitly to consumer welfare, and the remainder, given the economic theories of the time, were fully consistent with that value. The ten explicit bills were introduced by Representatives McRae (D., Arkansas); Fithian (D., Illinois); Henderson (R., Iowa); Conger (R., Iowa); Blanchard (D., Louisiana); Anderson (D., Mississippi); Enloe (D., Tennessee); Richardson (D., Tennessee); Lane (D., Illinois); and Perkins (R., Kansas).[36]

Explicit value statements in the Senate and the House, then were overwhelmingly in favor of the proposition that Congress intended the Sherman Act to be interpreted in accordance with the principles of consumer welfare. Those few legislators who spoke against that value were, as we shall see in a later section, in opposition to Sherman's bill and also, significantly, to the bill introduced by Senator Reagan (D., Texas).

II. The Proposed Rules of Law

I have already indicated the policy which underlies rules against cartel agreements (sometimes referred to as loose combinations), monopolistic mergers (tight combinations), and predatory tactics. In this section I will attempt to show that those rules were in fact contemplated by Congress.

a. Cartels

Doubt has been expressed about the clarity of the congressional intent with respect to cartels.[37] Yet it seems plain that Congress intended to outlaw "loose combinations" of the sort typified by price-fixing and market-division agreements between competitors. (I am speaking here of agreements not involving any significant efficiency-creating integration.) The evidence for this intent is of several sorts.

The language of the Sherman Act itself seems to distinguish between cartels and tighter arrangements similar to mergers. Section 1 refers to "Every contract, combination in the form of trust or otherwise, or conspiracy." Aside from the differing connotations of the words "contract," "combination," and "conspiracy," there is the obvious point that the drafters chose to modify only the word "com-

36 McRae, H.R. 91, Bills and Debates in Congress Relating to Trusts [hereinafter cited as Bills and Debates], S. Doc. No. 147, 57th Cong., 2d Sess. (1903), 417; Fithian, H.R. 202, *id.* at 421; Henderson, H.R. 270, *id*. at 425; Conger, H.R. 286, *id.* at 427; Blanchard, H.R. 402, *id*. at 431; Anderson, H.R. 509, *id*. at 433; Enloe, H.R. 811, *id*. at 435, and H. Rec. 30, *id*. at 459; Richardson, H.R. 826, *id*. at 437; Lane, H.R. 3819, *id*. at 449; and Perkins, H.R. 3844, *id*. at 451.

37 Thorelli, *op. cit. supra* note 8 at 185, seems to find Sherman's intention with respect to cartels, "simple agreements, pools and similar loose associations" somewhat ambiguous.

bination" with the phrase "in the form of trust or otherwise." The word "trust" originally gained currency to describe anticompetitive combinations because the trust device was used to gather industries or large parts of them under single ownership and control. It is arguable, therefore, that its use in the Sherman Act indicated that the "trust" was but one member of the general class of close-knit "combinations" while it was not a member of the classes of "contracts" or "conspiracies."[38] The distinction between the latter two terms may have been that between formal agreements and informal, probably secret, understandings. In any event, the obvious setting apart of the word "combination" in a way which seems to indicate common ownership and control suggests that something else was meant by the other two words, and that something else could hardly have been anything other than cartels. This argument is rather speculative, however, and clearer evidence exists.

Sherman's original draft of S.1, as well as Reagan's and the other bills, supports the theory that Congress intended to prohibit cartels by employing words that suggest every range of coordination from the loosest general understanding to the tightest-knit integration.[39] In the debates, moreover, Sherman plainly demonstrated an intention to outlaw cartels. He expounded his legislative aims, for instance, by reading to the Senate at great length from judicial opinions which he stated were representative of the common law he said his bill would enact. The cases he read from or described held illegal a market-division cartel agreement[40] as well as monopolistic mergers and the predatory extraction of railroad rebates by the Standard Oil Trust.

. . . SHERMAN PLAINLY DEMONSTRATED AN INTENTION TO OUTLAW CARTELS.

38 The word "trust" was used very loosely in the debates and sometimes, as in Reagan's bill, seemed to mean any arrangement that was formed for the purpose of suppressing competition. Some legislators, however, appeared to use the word to mean an arrangement involving integration by ownership. See Sherman's remarks at 21 Cong. Rec. 2457 (1890).

39 S.1 as drafted by Sherman applied to "all arrangements, contracts, agreements, trusts or combinations." The words suggest a progression from the loosest sort of understanding between independent firms to the tightest integration by ownership. George's bill, S.6, applied to "all contracts, arrangements, agencies, trusts, or combinations." Bills and Debates, 411. It is subject to similar analysis.

Reagan's amendment provided criminal penalties for persons creating or participating in trusts and specified that "a trust is a combination of capital, skill, or acts" for any or all of six stated purposes, among which were: creating or carrying out any restrictions in trade; limiting or reducing production or increasing or reducing prices; preventing competition; or creating a monopoly. Bills and Debates, 218. The use of the disjunctive seems significant, and a combination of either "skill" or "acts" alone could hardly be anything other than a cartel agreement. Coke's amendment in the Senate resembled Reagan's wording on this point, 21 Cong. Rec. 2613, (1890) as did a number of bills introduced in the House: H.R. 179, H.R. 830, H.R. 846, H.R. 3925, H.R. 8980. See Bills and Debates 419, 439, 441, 455, 457, respectively.

40 Chicago Gas Light and Coke Co. v. The People's Gas Light and Coke Co., 121 Ill. 530 (1887), held void on grounds of public policy an agreement dividing territories between two gas companies in Chicago. With respect to some of the other cases cited it is not clear whether the courts or Sherman viewed the arrangements as essentially mergers or cartels. See Craft v. McConoughy, 79 Ill. 346 (1875).

Sherman's intention to outlaw cartels was understood by his colleagues, and the remarks of Senators Stewart (R., Nevada) and Platt (R., Connecticut), who favored certain cartels show that not only Sherman's bill but Reagan's would enact such a rule. Stewart objected to both bills because they would ban competitors' agreements to limit output during periods of "overproduction" and "depression.[41] Platt attacked Sherman's measure because, "Unrestricted competition is brutal warfare . . . ."[42] He favored a rule that would permit agreements to charge prices that were "just and reasonable and fair."[43] The Senate paid no attention to either Platt or Stewart and, in the Committee of the Whole, adopted Reagan's amendment and reported Sherman's bill with its various additions and amendments to the Senate.[44]

The intention of both houses of Congress to outlaw cartels is also shown by the extended sparring that took place over the Bland amendment the House added to the Senate bill. Representative Bland (D., Missouri) offered a two-part amendment to make clear that the bill covered "every contract or agreement entered into for the purpose of preventing competition in the sale or purchase of any commodity, or to prevent competition in transportation."[45] The House had before it then the Senate Judiciary Committee's draft which may have seemed less clear than the Sherman and Reagan bills to those who had not followed the Senate debates. The House adopted Bland's amendment[46] but the Senate Judiciary Committee objected to the first part as beyond Congress' power under the commerce clause.[47] Indeed, the switch from Sherman's and Reagan's wording to the Judiciary Committee's seems originally to have been motivated largely by

41 21 Cong. Rec. 2605-2606; 2643 (1890). Stewart rather marred the consistency of his position when he read the price-fixing clause of Reagan's bill and announced: "If two or more persons fix the price at which they will sell any article they have got to go to the penitentiary. Well, I think they ought to. [laughter.]" *Id.* at 2644.

42 21 Cong. Rec. 2729 (1890).

43 *Ibid.* Platt said, "The conduct of this Senate for the past three days . . . has not been in the line of the honest preparation of a bill to prohibit and punish trusts. It has been in the line of getting some bill with that title that we might go to the country with. . . . [T]he whole effort has been to get some bill headed 'A bill to punish trusts' with which to go to the country." *Id.* at 2731. This remark is often quoted to suggest that Congress had no serious intention in passing the Sherman Act. More likely, however, the statement simply reflects Platt's strong disapproval of a measure which would outlaw cartels he thought desirable. Platt was opposed to Sherman's consumer-welfare policy, and his charge should be evaluated with that in mind.

44 Reagan's amendment was adopted by a vote of 34 to 12 on March 25, 1890. 21 Cong. Rec. 2611 (1890). S.1 was reported from the Committee of the Whole to the Senate on March 26. *Id*. at 2662.

45 *Id*. at 4099.

46 *Id.* at 4104.

47 *Id.* at 4559-4560.

the very doubts of constitutionality which Bland's amendment provoked. Senator Hoar (R., Massachusetts), who reported the Judiciary Committee's reaction to the Bland amendment stated he thought the remainder of the amendment concerning transportation was covered in the Senate bill already, but there was no harm in adding Bland's proposal.[48] Had matters stopped there it would have been clear that cartels were illegal. If the Senate bill, which became the Sherman Act, covered railroad rate cartels (which was what the House was driving at in the second part of the Bland amendment), it certainly covered other cartels. But for the question of the reach of the commerce power, the first part of the Bland amendment was surely covered by a bill which made no distinction between transportation and other goods or services.

A day later, however, Hoar said some members thought the Judiciary Committee's revision of the Bland amendment was not as precise and well guarded as it might be. He did not explain, but his motion to recommit was agreed to.[49] The committee came back with a very different amendment under which agreements preventing competition in transportation were illegal only if rates were "raised above what is just and reasonable."[50] The Senate agreed,[51] perhaps because this approach to transportation seemed more in keeping with the railroad rate philosophy of the recently enacted Interstate Commerce Act. The House refused to accept the Senate amendment and ultimately both the Senate and the House agreed to recede from their respective amendments, leaving the bill as it had first come from the Senate.[52]

The inference from this maneuvering is that all cartels were to be illegal, regardless of the price they set. The Senate Judiciary Committee's objection to the first part of the Bland amendment is probably to be taken at face value. Hoar's statement that the bill already covered the second part of the Bland amendment indicates not only that the unamended bill made railroad cartels flatly unlawful but that it had that effect upon all other cartels since there is nothing in the wording of the statute or the debates to suggest the Senate had intended a distinction. Indeed, Hoar's words may have been the factor that galvanized the senators who favored a different rule for railroads to press for a revision specifying a "just and reasonable" standard for railroad rate agreements. This move constitutes an admission that the general language of the bill permitted no such construction. By receding afterward the Senate appears to have indicated again that the flat rule applied to all cartels.

48 *Id*. at 4560.

49 *Id*. at 4599.

50 *Id*. at 4735.

51 *Ibid.*

52 *Id.* at 5950-5961; 5981-5983; 6116-6117; 6312-6314.

No contrary implication can be drawn from the House's recession. The House had not attempted to distinguish between railroad and other cartels. By the time of the recession, particularly in view of Hoar's first statement, it may very well have seemed that Bland's amendment was unnecessary to the House's purposes.[53]

The evidence appears unmistakable that the Congress intended to outlaw cartels.

b. and c. Monopolistic mergers and predatory practices

There is no need to spell out all the evidence that Congress intended to outlaw both mergers (or other forms of close-knit combination) that created monopoly and predatory business tactics. Sherman's description of the common law which his bill would enact,[54] his other remarks,[55] the speeches of a number of legislators,[56] and the language of the bills introduced[57] sufficiently establish this point. No one, to my knowledge, has ever challenged it. The important point is that these rules were typically justified in terms of consumer welfare. Sherman stated the general case against both monopolistic mergers and predatory practices:

---

53 See the remarks of Culberson, *id.* at 5951 (as to the first part of Bland's amendment), and of Edmunds, *id.* at 6116; Vest, *ibid*. and *id.* at 4123; and Hoar (as to the second part of Bland's amendment), *id*. at 4560.

54 Handy v. Cleveland and Marietta Railroad Co., 31 Fed. 689 (1887), involved the predatory extraction of railroad rebates by the Standard Oil Co.; Richardson v. Buhl, 77 Mich. 632 (1889); People v. Chicago Gas Trust Co., 130 Ill. 268 (1889); and People v. North River Sugar Refining Co., 121 N.Y. 582, 623, 626 (1890), involved monopolistic mergers and acquisitions.

55 See 21 Cong. Rec. 2457, 2459-2462, 2569 (1890).

56 *E.g*., Edmunds, *id.* at 2726; Reagan, *id.* at 2645; Pugh, *id*. at 2558; Culberson, *id.* at 4089.

57 Reagan's amendment clearly aimed at predatory practices and monopolistic mergers as well as cartels. Among the purposes for which it was forbidden to enter into "a combination of capital, skill, or acts" was "to limit or reduce the production or to increase or reduce the price of merchandise or commodities." Bills and Debates, 218. This provision seems curiously asymmetrical at first glance. Reduction of output and increase of prices occur together. The seemingly anomaly of permitting combinations to increase production but forbidding combinations to lower prices may, however, be resolved by the theory of predation. A combination formed with the intention of increasing production may have seemed to Reagan to display an intention to create efficiency by cutting costs. But the intention to lower prices may have seemed unrelated to costs and therefore to imply a further intent to injure rivals improperly. This interpretation is given substance by the fact that Reagan introduced much the same bill as S.3440 in the 50th Congress, but there listed as a prohibited purpose "To limit, to reduce, or to increase the production or prices of merchandise or commodities." Bills and Debates, 5. His later bill in the same Congress, S. 3476, amended this so as not to prohibit an intent to increase production. Bills and Debates, 33, and his amendment to S.1 in the 51st Congress followed that pattern. The parallel bills offered in the House by Anderson, H.R. 11213 and H.R. 11279, underwent the same evolution. Bills and Debates, 55 and 57. These changes strongly suggest an attempt to preserve efficiency-creating combinations while prohibiting predatory combinations.

Reagan's bill also prohibited combinations to create monopolies. It has already been shown, note 39, *supra*, that it prohibited cartels. Its adoption by the Senate in Committee of the Whole, therefore, strongly supports the argument that the Senate proposed to enact the rules discussed in the text.

> "The sole object of such a combination is to make competition impossible. It can control the market, raise or lower prices, as will best promote its selfish interests, reduce prices in a particular locality and break down competition and advance prices at will where competition does not exist. Its governing motive is to increase the profits of the parties composing it. The law of selfishness, uncontrolled by competition, compels it to disregard the interest of the consumer. It dictates terms to transportation companies, it commands the price of labor without fear of strikes, for in its field it allows no competitors. Such a combination is far more dangerous than any heretofore invented, and, when it embraces the great body of all the corporations engaged in a particular industry in all the States of the Union, it tends to advance the price of the consumer of any article produced, it is a substantial monopoly injurious to the public, and, by the rule of both the common and the civil law, is null and void and just subject of restraint by the courts, of forfeiture of corporate rights and privileges, and in some cases should be denounced as a crime, and the individuals engaged in it should be punished as criminals. It is this kind of a combination we have to deal with now."[58]

The emphasis in this passage is upon the harm done to consumers. Sherman also mentions that a combination of the sort he describes allows no strikes but that is clearly an additional evil and not a test for illegality to be applied independently of consumer welfare. If there were any ambiguity in the passage, it would be removed by the wording of his bill which specifies only consumer-welfare tests. Other legislators spoke of the evils of the trusts in respects other than their harmful effect upon consumers, but, like Sherman, none of them suggested that these harmful effects could take place in any case not involving injury to consumers. The language is always fully consistent with the view that concern for farmers, laborers, or small businessmen was complementary to concern for consumers and not to override it in case of conflict between the interests of consumers and other groups. Other factors, particularly the one to be discussed in the next section, demonstrate that this interpretation is the correct one.

58 21 Cong, Rec. 2457 (1890).

III. The Preservation of Efficiency; the Legality of Monopoly Gained Through Efficiency

Congress' position with respect to efficiency cannot be explained on any hypothesis other than that consumer welfare was in all cases the controlling value under the Sherman Act.

Sherman took great pains to stress that his bill would in no way interfere with efficiency. It would outlaw only those mergers which created great market power. "[The bill] aims only at unlawful combinations. It does not in the least affect combinations in aid of production where there is free and fair competition."[59] He stressed the legality of efficiency repeatedly,[60] citing partnerships and corporations as two forms of combination which were efficiency-creating and therefore lawful.[61] He said corporations "ought to be encouraged and protected as tending to cheapen the cost of production."[62] He also praised the efficiency-creating corporate merger.[63]

---

59 *Ibid.* The words "in aid of production" obviously refer to efficiency.

60 *E.g.,* "If their [the individuals'] business is lawful they can combine in any way and enjoy the advantage of their united skill and capital, provided they do not combine to prevent competition." *Ibid.*

61 Sherman stated the case for partnerships:

> The right to combine the capital and labor of two or more persons in a given pursuit with a community of profit and loss under the name of a partnership is open to all and is not an infringement of industrial liberty, but is an aid to production . . . . The same business is open to every other partnership, and, while it is a combination, it does not in the slightest degree prevent competition.

Ibid.

Sherman attributed to the corporate form of combination an efficiency-creating potential even greater that that of partnerships, presumably because corporate enterprise is likely to operate on a larger scale. In any event, Sherman's praise for the corporate form of organization strikingly demonstrated that he wished to preserve efficiency precisely because it enriches consumers:

> The combination of labor and capital in the form of a corporation to carry on any lawful business is a proper and useful expedient, especially for great enterprises of a quasi public character, and ought to be encouraged and protected as tending to cheapen the cost of production, but these corporate rights should be open to all upon the same terms and condition. . . . Experience has shown that they are the most useful agencies of modern civilization. They have enabled individuals to unite to undertake great enterprises only attempted in former times by powerful governments. The good results of corporate power are shown in the vast development of our railroads and the enormous increase of business and production of all kinds.

Ibid.

62 See the last quotation in note 61, *supra*.

63
> When corporations unite merely to extend their business, as connecting lines of a railway without interfering with competing lines, they are proper and lawful. Corporations

Not once did Sherman suggest that courts should blunt or discourage efficient size or conduct in the interest of any social or political value. The only limit he urged to the creation of efficiency by combination was justified explicitly in terms of consumer welfare. He thought combinations of monopolistic size would not pass their efficiencies on to consumers:

> "It is sometimes said of these combinations [the monopolistic trusts] that they reduce prices to the consumer by better methods of production, but all experience shows that this saving of cost goes to the pockets of the producer. The price to the consumer depends upon the supply, which can be reduced at pleasure by the combination."[64]

Here again Sherman identified injury to consumers as occurring through restriction of output by firms with market control.

The Senate later adopted an amendment to Sherman's bill offered by Senator Aldrich (R., Rhode Island) which stated:

> "*Provided further*, That this act shall not be construed to apply to or to declare unlawful combinations or associations made with a view or which

footnote 63 cont'd

> tend to cheapen transportation, lessen the cost of production, and bring within the reach of millions comforts and luxuries formerly enjoyed by thousands.

21 Cong. Rec. 2457 (1890).

The progression in Sherman's argument-partnerships to corporations to corporate mergers-indicates that he perceived these three forms of integration as essentially the same economic phenomenon. All are capable of benefiting consumers by creating efficiency.

64 *Id.* at 2460. Sherman's argument here is not necessarily correct. A monopolistic merger may create such efficiency that the net effect will be an increase in output. There is no way of telling in advance, or even afterward, in all probability, whether the net effect of such a merger will be restriction or increase of output.

The passage quoted in notes 61 and 63 *supra*, and in the text here show that when Sherman proposed to outlaw the prevention of "full and free competition" he did not refer to any elimination of rivalry between firms. The combinations he favored all eliminate such rivalry. The phrase "full and free competition" must be read to refer to a market whose structure is effectively competitive. Sherman gives few clues as to the structure he envisaged, though he seemed willing to allow rather high percentages. It is not necessary to conclude that Sherman would have required a market power test for cartels because such agreements, not being "in aid of production," could not benefit consumers and could only be motivated by a desire to restrict output.

> tend, by means other than by a reduction of the wages of labor, to lessen the cost of production or reduce the price of any of the necessaries of life, nor to the combinations or associations made with a view or which tend to increase the earnings of persons engaged in any useful employment."[65]

The adoption of this amendment by the Senate in Committee of the Whole indicates agreement with Sherman's position on efficiency, though Thorelli warns that at the time of adoption the Senate was concerned primarily with Ingalls' amendment to prohibit trading in futures and options.[66] The last clause of the amendment appears merely to reflect the Senate's desire to exempt labor unions from the scope of the law.

The most dramatic illustration of Congress' agreement with Sherman's position, however, was the decision to make legal the gaining of monopoly by superior efficiency. The Judiciary Committee draft made it an offense to "monopolize," not to have a monopoly. The wording itself suggests that an activity rather than a status was to be outlawed, and that in turn suggests that there were lawful means of gaining a monopoly position. The issue was raised by Senator Kenna (D., West Virginia) who asked:

> "Is it intended by the committee, as the sections seems to indicate, that if an individual engaged in trade . . . by his own skill and energy, by the propriety of his conduct generally, shall pursue his calling in such a way as to monopolize a trade, his action shall be a crime under this proposed act?"

65 21 Cong. Rec. 2654-2655 (1890).

66 Thorelli, *op. cit. supra* note 8, at 195. Aldrich's suggestion, however, accords with Sherman's position and it seems reasonable that the Senate agreed with his amendment when it adopted it. No senator spoke against efficiency and others praised it. Teller, for example, pointed out that "A trust may not be always an evil. A trust for certain purposes, which may mean simply a combination of capital, may be a valuable thing to the community and the country." 21 Cong. Rec. 2471 (1890). Blair was seemingly concerned about efficiency when he suggested that Sherman's bill be amended by striking out the words "to prevent full and free competition" and inserting in their place the words "to permit a monopoly," and also to change the phrase "intended to advance the cost" to read "primarily intended to enhance [sic]" the cost to the consumer. *Id.* at 2566-2567. Blair was troubled by the ambiguity of the word "competition" and probably wished by the first change to ensure that the courts did not strike down every combination that eliminated some rivalry. The second suggestion seems to recognize that combinations may be foreseen to have both an output-restricting and an efficiency-creating tendency and that only those made with the primary intent of restricting output should be unlawful. The court would have to weigh or assess which of the two contradictory tendencies was intended or expected to predominate in order to predict the net impact for good or ill upon consumers. Reagan's apparent desire to preserve efficiency-creating combinations is discussed in note 57, *supra.*

Kenna then framed a hypothetical:

> "Suppose a citizen of Kentucky is dealing in shorthorn cattle and by virtue of his superior skill in that product it turns out that he is the only one in the United States to whom an order comes from Mexico for cattle of that stock for a considerable period, so that he is conceded to have a monopoly of that trade with Mexico; is it intended by the committee that the bill shall make that man a culprit?"[67]

This example was somewhat muddy since it was not clear whether the trader had achieved control of the supply of cattle or had merely had success with Mexican customers. Edmunds' initial answer to the question preserved the ambiguity,[68] but further discussion clarified the committee's intent.

Senator Hoar said he had put in the Judiciary Committee the precise question asked by Kenna because he had the same difficulty. He was answered, and he thought all the members of the committee agreed to the answer, that "monopoly" was a technical term known to the common law. The "clear and legal signification" of the term, said Hoar, showed: "It is the sole engrossing to a man's self by means which prevent other men from engaging in fair competition with him." Hoar then went on to remove the ambiguity from Kenna's hypothetical:

> "I suppose, therefore, that the courts of the United States would say in the case put by the Senator from West Virginia that a man who merely by superior skill and intelligence, a breeder of horses or raiser of cattle, or manufacturer or artisan of any kind, got the whole business because nobody could do

67 21 Cong. Rec. 3151 (1890).

68 Kenna had gone on to say that the bill provided a penalty for any citizen "who happens by his skill and energy to command an innocent and legitimate monopoly of a business." Edmunds replied, "It does not do anything of the kind, because in the case stated the gentleman has not any monopoly at all. He has not bought off his adversaries. He has not got the possession of all the horned cattle in the United States. He has not done anything but compete with his adversaries in trade, if he had any, to furnish the commodity for the lowest price." *Id*. at 3151-3152. Since Edmunds mentioned both the absence of possession of all the cattle and the absence of merger or improper tactics, it would be difficult from this answer alone to say that real monopoly gained through efficiency was intended to be lawful.

Gray proposed to cure the difficulty Kenna pointed out by amending section 2 to require a combination or conspiracy to monopolize. *Id*. at 3152. This would have left a gap in the law for single-firm conduct and so was rejected. It indicates, however, that Gray did not wish growth by efficiency hindered in any way.

> it as well as he could was not a monopolist, but that it involved something like the use of means which made it impossible for other persons to engage in fair competition, like the engrossing, the buying up of all other persons engaged in the same business."[69]

Hoar's answer, then, is that monopolies gained by merger or predatory tactics are illegal but monopolies gained by superior efficiency are not. Edmunds then explained further and supported Hoar's position.[70] No contrary position with respect to the desirability of legalizing monopoly gained through efficiency or the meaning of the statute was expressed. Apparently satisfied with the construction put on Section 2 by Hoar and Edmunds, the Senate promptly passed the bill, 52 to 1.

> THE SENATE'S CONSCIOUS ELECTION TO LEGALIZE MONOPOLY BY EFFICIENCY, THEREFORE, IS HIGHLY SIGNIFICANT—A CLEAR CHOICE OF CONSUMER WELFARE AND THOSE VALUES CONSISTENT WITH IT OVER COMPETING VALUES . . .

Congress' decision to permit monopoly achieved by efficiency is completely inconsistent with the view that courts should use the Sherman Act to ameliorate the noneconomic "helplessness of the individual" before "great aggregations of capital" or that they may take into account the alleged desirability of preserving for its own sake an economy of small business units. Monopoly by efficiency is as effective as monopoly by predation in driving smaller rivals from an industry, and it would seem to have whatever undesirable social or political side effects that any monopoly or large industrial size may be thought to imply. Monopoly by

69 *Ibid.*

70 Edmunds said the best answer he could give to Kenna was to read from Webster's Dictionary the definition of the verb "to monopolize":

> 1. To purchase or obtain possession of the whole of, as a commodity or goods in market, with the view to appropriate or control the exclusive sale of; as, to monopolize sugar or tea.

Edmunds interjected: "Like the sugar trust. One man, if he had capital enough, could do it just as well as two." He went on from the dictionary:

> 2. To engross or obtain by any means the exclusive right of, especially the right of trading to any place, or with any country or district; as, to monopolize the India or Levant trade.

Id. at 3152.

The second definition quoted may make Edmund's first answer to Kenna, note 68, *supra*, more meaningful. Kenna's Kentuckian had gotten all the trade with a country, and Edmunds, though his dictionary indicated that such trade could be "monopolized," had stated that under Kenna's hypothetical circumstances the cattle trader had not violated the statute. Perhaps Edmunds' two answers may be made consistent if the first answer is read as relying upon the absence of merger and predatory tactics rather than upon lack of possession of all the cattle in the United States.

efficiency, however, is probably beneficial to consumers and to small business suppliers and customers of the monopolists—at least by comparison with the policy alternative. Breaking up monopoly gained by efficiency is likely to impose higher costs at that level of the distributive or productive chain to the detriment of consumers and all vertically related firms. The Senate's conscious election to legalize monopoly by efficiency, therefore, is highly significant—a clear choice of consumer welfare and those values consistent with it over competing values, including that of preserving small business units in the same market.

IV. Proposals to Exempt Labor and Farm Organizations

A number of senators spoke in favor of exempting from the statute's coverage organizations of laborers to raise wages and organizations of farmers to raise the price of farm products.[71] The Senate in Committee of the Whole adopted such exemptions,[72] but Edmunds opposed them,[73] and when the bill came back from the Judiciary Committee, where Edmunds had played a major part in its phrasing, no explicit exemption remained. It may be debatable, since some senators had thought the exemption inherent in the bill without being expressed, whether the Senate intended the exemption or not. The significant fact for present purposes, however, is that not one legislator suggested that the conflicting values of consumer interest versus farmer and laborer interests be delegated to the courts for resolution case by case. The universally favored techniques were either full exemption from or full application of the statute. The Senate's all-or-nothing approach here, where many clearly regarded conflicting values as in play, tends to buttress the view that all cases to which the statute did extend were to be decided exclusively upon considerations related to consumer welfare.

V. The Narrow Scope of the Commerce Power

The notion abroad today that the Fifty-first Congress breathed broad social and political values into the Sherman Act is an anachronism. The Congress and the Supreme Court of 1890 had no such expansive view of federal power generally, and of the commerce power in particular, as has become familiar in recent times. The limitations upon Congress' commerce power were thought to be of two

71 Sherman, 21 Cong. Rec. 2562, 2611 (1890); George, S.6, Bills and Debates, 411; Teller (objecting to Sherman's and Reagan's bills for not having exemptions for labor and farm organizations), 21 Cong. Rec. 2561-2562 (1890); Hoar, *id.* at 2728; Coke, *id*. at 2615; Gray (offering George's bill as an amendment to S.1), *id.* at 2657; Aldrich, *id*. at 2654-2655; Hiscock (objecting to the lack of a labor exemption in Sherman's bill), *id*. at 2468; Stewart, *id.* at 2643.

72 *Id.* at 2612.

73 *Id.* at 2726-2729.

sorts—the reach of the power, defined by the interstate-intrastate distinction, and the nature of the power, defined by the commercial-noncommercial distinction. These limitations are related, both being based on concepts of federalism and limited central government. For that reason it is logically and psychologically probable that men who favored a short reach in the commerce power would favor a narrow definition of the goals for which the power could be exercised. There is ample evidence in the Congressional Record that the Fifty-first Congress took a limited view of the reach of the interstate concept[74] and it is correspondingly unlikely that they took a broad view of the values the power could be used to implement directly.

More direct evidence of Congress' view of the goals to be directly implemented through the commerce power comes both from the general trend of legislation under that clause and statements made in the course of the passage of the Sherman Act. The first major commerce clause legislation was the Interstate Commerce Act of 1887. Congress there confined its law to matters bearing directly upon the movement of commerce—terms and conditions of interstate transportation by railroad. It apparently believed that the commerce power did not enable Congress to prevent the starting of unnecessary railroad enterprises or to regulate railroad financial operations such as fictitious capitalization.[75] Up to 1890 Congress had not even attempted to exercise a general "police power" under the commerce clause, and it was not until 1895 that a very modest beginning was made with a statute barring lottery tickets from movement in interstate commerce. In 1903, the Supreme Court, divided five to four, upheld the statute as within the commerce power. Even then the majority felt obliged to use a "pollution-of-commerce" rationale, analogizing the statute to the prohibition of the interstate movement of diseased cattle.[76] Congress moved slowly into the field of social legislation, and the Supreme Court struck many such laws down for over forty years after the passage of the Sherman Act. Since Congress was experimenting timidly with the commerce power as a vehicle for social reform well after 1890, and the Supreme Court was resisting well into the 1930's. It seems farfetched to suppose that Congress intended to enact broad social welfare measures through the Sherman Act.

This general argument is borne out by the legislative history of the statute. Sherman's argument for the constitutionality of his bill rested entirely on the theory that it would facilitate the flow of interstate commerce: "[Congress] may 'regulate commerce;' can it not protect commerce, nullify contracts that restrain

74 *E.g.,* Hoar, *id*. at 2568; Reagan, *id*. at 2469-2470, 2601; Stewart, *id*. at 2566; Gray, *id*. at 2657; Coke, *id*. at 2614.

75 Elder, A Handbook of the Interstate Commerce Act 4 (1931).

76 Champion v. Ames, 188 U.S. 321 (1903).

commerce, turn it from its natural courses, increase the price of articles, and therefore diminish the amount of commerce?"[77]

Edmunds, who proposed the final phrasing of the statute's relation to commerce in the Judiciary Committee, took a very limited view of federal power. He said the Constitution did not give and ought not to give Congress "power to enter into the police regulations of the people of the United States."[78] Edmunds maintained that Congress lacked the power under the commerce clause to abolish the sugar trust.[79] He opposed Ingalls' proposal to tax dealings in options and futures because it was essentially a "police measure" and the Supreme Court would say that Congress had no power "to regulate the good order of society."[80] It is hardly conceivable that a man with such views could have drafted a bill intended to hand over to federal courts, operating under a delegation of the commerce power, the right to adjust social and political ills of a noncommercial nature.

The Judiciary Committee dropped the wording of Sherman's and Reagan's bills and instead employed the phraseology not merely of the common law but of Sherman's reasoning about Congress' power over commerce. The redrafted bill spoke in terms of the diminution or lessening of the flow of commerce which Sherman had said resulted from control of the market—that is, in terms of contracts, combinations and conspiracies "in restraint of trade or commerce" and monopolizations of such trade or commerce. This adroit phrasing not only imported the substantive criteria which Sherman had proposed but was calculated to satisfy both broad and narrow constructionists of the commerce power's reach. The Act's reach would depend upon the Supreme Court's demarcation of the line between interstate and intrastate. But the wording also indicated that the test for illegality was entirely the effect upon commerce, not an effect upon some other thing or condition, such as a supposed social or political evil, which had merely some requisite jurisdictional effect upon commerce.

---

77 21 Cong. Rec. 2462 (1890). This reading of Sherman is confirmed by his subsequent remark:

> [T]he object aimed at by this bill is to secure competition of the productions of different States which necessarily enter into interstate and foreign commerce. These combinations strike directly at the commerce over which Congress alone has jurisdiction. "Congress may regulate interstate and foreign commerce," and it is absurd to contend that Congress may not prohibit contracts and arrangements that are hostile to such commerce.

Ibid.

The phrases "strike directly" at commerce and "hostile to" commerce could hardly be employed to mean anything other than diminish commerce.

78 *Id.* at 2727.

79 *Id.* at 2728.

80 *Ibid.*

This evidence of the Fifty-first Congress' view of the scope of the commerce power is of course not conclusive of the point sought to be established here. But it does tend strongly to indicate the improbability of the proposition that Congress intended to delegate noncommercial criteria to the federal courts. More than that, Sherman's commerce clause argument and the wording of the final bill suggest not merely that the statute's intended goals were commercial but that they related entirely to safeguarding the flow of commerce against diminution, against, in a word, a restriction of output.[81] Edmunds' views on this point were shared by other members of the Judiciary Committee and the Senate.[82] Edmunds told the Senate that the Judiciary Committee had unani-

---

81 See note 77 *supra* and related text. It is significant that the Senate did not discuss its power "to regulate the good order of society" in connection with Sherman's or Reagan's bills, or any other antitrust proposals, but only with relation to Ingalls' proposed prohibition of dealings in options and futures. This strongly suggests that the antitrust proposals were not viewed as attempts to reach values or ends other than the removal of obstructions to the free flow of interstate commerce.

82 Seven of the nine members of the Judiciary Committee expressed views which indicate their conceptions of the scope of Congress' power under the commerce clause. Six of the seven appear to have held views in 1890 which indicate they would not have supported any bill designed to accomplish social or political objectives unrelated to the freedom and volume of the flow of interstate commerce. Such objectives many of them thought entirely reserved to the "police power" of the individual States. These senators were Edmunds, the committee chairman, Hoar, Vest, George, Coke, and Pugh. The seventh, Ingalls, while he did not speak of the commerce clause, expressed such a broad view of constitutional powers generally that he might well have been willing to accept a federal "police power" under the commerce clause.

Edmunds views have been cited in the text. Hoar thought that Congress' only jurisdiction was to "protect" interstate and foreign commerce, 21 Cong. Rec. 2567 (1890). He apparently thought the commerce power limited to freeing the flow of commerce and to transactions rather immediately concerned with an actual interstate movement.

Vest was more explicit. He said the Constitution did not give Congress power "to legislate as it sees proper, under the general and nebulous presumption of the general welfare." *Id.* at 2463. He seemed to think that the power did not extend beyond a strict definition of commerce. *Id*. at 2465.Vest thought that even Sherman's bill was unconstitutional and said it "destroys all my ideas of the limitations of the Constitution." *Id.* at 2570. He favored Coke's bill which merely prohibited shipment of trust products out of any State that had declared trusts unlawful. *Id.* at 2570-2571. He insisted that "the police power of the State is an entirely different jurisdiction, as distinct and separate from the interstate-commerce clause in the Federal Constitution as any two subjects can possibly be." *Id.* at 2603. See also, *Id.* at 2645 and 4560.

George seemed to take a more limited view of the commerce power than any other senator. He thought, as did others, that the ends to be served by legislation under the commerce power had to be commercial in nature. *Id*. at 1768, 1770. He thought that even Sherman's bill was far too broad. *Id*. at 1771. Indeed, George said Congress could not reach the prevention of full and free competition in manufacturing because there was nothing in such a rationale that "would not authorize Congress to make any other regulation they might deem wise in such production . . . ." *Id.* at 1769. He was clear that there was no such power under the commerce clause and that nothing was added by the fact that the manufacturers competed with imported goods on which a duty was paid. See generally, *id.* at 1768-1772. Reagan's bill he thought also unconstitutional. *Id*. at 2560. In fact, at one point, George thought no conceivable worthwhile legislation could be laid under the commerce power. *Id.* at 2598. See also, *id*. at 2598-2600. He thought even the draft of the Judiciary Committee would prove disappointing to the people. "It covers professedly a very narrow territory, leaving a very large number of these institutions, these trusts, or whatever we may call them, entirely without the purview of the bill. That is not mously determined to "frame a bill that should be clearly within our constitutional power."[83] What discussion there was on the topic in the House of Representatives paralleled the majority position in the Senate.[84]

---

footnote 82 cont'd

the fault of the Committee, Mr. President. The bill has been very ingeniously and properly drawn to cover every case which comes within what is called the commercial power of Congress." *Id.* at 3147.

Coke as has been noted, drafted a bill which backed up State police power by prohibiting transportation of goods made by trusts from States that declared such trusts unlawful. The only other sanction was a presidential power and duty to suspend duty collections on goods of the type controlled by a domestic trust. He thought Sherman's and Reagan's bills liable to George's constitutional objections. *Id.* at 2614. Coke's belief that Congress could do little more than support State police power suggests that he thought Congress had no police power under the commerce clause which would enable it to reach noncommercial ends.

Pugh defended the constitutionality of Sherman's bill on a rationale much like Sherman's. That is, Pugh argued that the agreements and combinations described in the bill "hinder, interrupt, and impair the freedom and fairness" of commerce. *Id.* at 2558. The only arguable word is "fairness" but Pugh was clearly not referring to a moral standard independent of the free flow of commerce. Sherman's bill, which Pugh quoted, would bear no such construction. Probably the word refers to predatory practices which were thought to impede the free flow of commerce, or possibly, to the unfairness to consumers with which Sherman's bill was explicitly concerned.

The other three members of the Judiciary Committee said nothing directly in point. It may be conceded, for the sake of argument, that Ingalls' disquisition on the tax power, in connection with his bill on dealings in options and futures, reveals a frame of mind which was not likely to find difficulties in achieving any desired end though an exercise of any power, though it seems a trifle odd that he did not attempt to lay his bill under the commerce power. See *id*. at 2648-2652. Wilson's remarks on the commerce power are beside the present point, *id.* at 2602-2604, except for an offhand remark that the police powers belong wholly to the States, *id.* at 2604. Evarts appears to have said nothing at all on the subject.

Among the senators not on the Judiciary Committee who spoke to this point, a similar majority took a restricted view of the commerce power which rules out any intention to accomplish noneconomic objectives through the Sherman Act. Sherman, as we have seen, invoked the commerce power on purely economic grounds. Even Reagan thought that the primary attack upon trusts had to be made by the States and that Sherman's bill went beyond the commerce power. *Id*. at 2469-2470, 2601. His view of the reach of his own bill is shown by his remark that, since it rested on the commerce power, farm and labor organizations would probably not be affected by it. *Id.* at 2561-2562. Eustis, in opposing Ingalls' amendment concerning futures and options, contended that Congress had no power to reach "the whole question of police, of policy, and of public morality." *Id*. at 2646. See also, *id*. at 2651-2652. Turpie, on the other hand, may have believed that Congress had a police power with respect to interstate commerce as broad as that of the States with respect to intrastate commerce. *Id.* at 2557.

83 *Id.* at 3148.

84 In reporting the bill Culberson assured the House:

> There is no attempt to exercise any doubtful authority on this subject, but the bill is confined strictly and alone to subjects over which, confessedly, there is no question about the legislative power of Congress . . . .

Id. at 4089.

VI. The Criteria Delegated to the Federal Courts

Those who have described the Sherman Act as the delegation of broad discretion to the courts, in some respects comparable to the powers delegated to or assumed by the courts under the great clauses of the Constitution, are of course quite correct. Congress specified a value, a core of meaning, and left it to the courts to elaborate a framework of subsidiary rules in the course of examining great numbers of market structures and forms of market behavior over a period of many years. But those who, like Judge Hand, think the delegation essentially unconfined are in error. Many legislators in the Fifty-first Congress remarked the fact of delegation, but none suggested that it was without standards.[85] The standards intended can easily be found.

As always John Sherman provides the clearest and best statement on the subject. Speaking of his own bill, whose policy, as we shall see, is to be equated with that of the subsequent Judiciary Committee draft that became law, Sherman said:

> "The first section, being a remedial statute, would be construed liberally, with a view to promote its object. It defines a civil remedy, and the courts will construe it liberally; they will prescribe the precise limits of the constitutional power of the Government; *they will distinguish between lawful combinations in aid of production and unlawful combinations to prevent competition and in restraint of trade . . . .*"[86] (Emphasis added.)

Sherman could hardly have said more clearly that the law was to delegate to the courts the task of distinguishing between those arrangements and combinations which increase efficiency and those that restrict output.

---

85 Aside from Sherman, Edmunds and Turpie in the Senate referred specifically to the fact that the statute would delegate much to the courts. *Id*. at 3148 and 2558. In the House Culberson, Wilson, Bland, Cannon, Morse, and Kerr referred to the delegation. *Id*. at 4089, 4092, 4099 and 5953, 4099, 5953, and 6313, respectively. The House had not of course heard the Senate debates which gave content to the bill's words, and several of the speakers there viewed the delegation as dangerously vague.

86 *Id.* at 2456. Later in the same speech, his main presentation of the topic to the Senate, Sherman stated: "I admit that it is difficult to define in legal language the precise line between lawful and unlawful combinations. This must be left for the courts to determine in each particular case. All that we, as lawmakers, can do is to declare general principles, and we can be assured that the courts will apply them so as to carry out the meaning of the law, as the courts of England and the United States have done for centuries. This bill is only an honest effort to declare a rule of action . . . ." *Id.* at 2460. The declaration of a "rule of action" is hardly equivalent to the bestowal of unconfined discretion. The reference in this passage to the common law also defines the range of the courts' discretion, as will be shown in the text.

Judge Hand was correct, of course, in viewing the Sherman Act's common law terminology as expressing the delegation of discretionary powers to the courts, but he and many other commentators appear to have misinterpreted the role of "the common law" in Sherman Act adjudication. The problem seems at first more difficult than it is because there was in 1890 no unitary body of common law doctrine which could give meaning to the statute. The common law of restraints of trade and monopolies has been a variable growth, composed of diverse and even contradictory strains, many of them obviously irrelevant to the concerns of the Sherman Act. Yet Sherman and many of his colleagues repeatedly assured the Senate, without objection by anyone, that they proposed merely to enact the common law.

There is no mystery, for Sherman and the others also repeatedly stated what the common law was. The fact that their statements did not accurately mirror that confused body of precedent does not obscure what they intended to convey. It is clear from the debates that "the common law" relevant to the Sherman Act is an artificial construct, made up for the occasion out of a careful selection of recent decisions from a variety of jurisdictions plus a liberal admixture of the senators' own policy prescriptions. It is to this "common law," holding full sway nowhere but in the debates of the Fifty-first Congress, that one must look to understand the Sherman Act.

I have already mentioned that the only cases cited by Sherman as representative of the common law held illegal and predatory extraction of railroad rebates by the Standard Oil Co., cartel agreements, and monopolistic mergers. But this extensive discussion of "the" common law was by no means the only occasion upon which Sherman told the Senate what the law was. He identified his bill—which struck at agreements preventing full and free competition or tending to advance costs to consumers—with the common law. The first point in Sherman's first speech on behalf of his bill was the categorical assertion that the bill "does not announce a new principle of law, but applies old and well recognized principles of the common law to the complicated jurisdiction of our State and Federal Government."[87] And later: "It is the unlawful combination, tested by the rules of common law and human experience, that is aimed at by this bill, and not the lawful and useful combinations."[88]

Sherman defined "monopoly" with a quotation from one of his selected common law cases: "Any combination the tendency of which is to prevent competition in its broad and general sense, and to control, and thus at will enhance, prices to the detriment of the public, is a legal monopoly."[89] And he concluded

87 *Id.* at 2456.

88 *Id.* at 2457.

89 *Id.* at 2459.

his review of the decisions by claiming for the common law generally a policy uniformity despite a variability in the law itself:

> "I might add to the cases cited innumerable cases in nearly all the States and in England, and in all of them it will appear that while the law in respect to contracts in restraint of trade and combinations to prevent competition and to advance the price of necessaries of life has varied somewhat, but in all of them, whether the combinations are by individuals, partnerships, or corporations, when the purpose of the combination or its plain tendency is to prevent competition, the courts have enforced the rule of the common law and have vigorously used the judicial power in subverting them."[90]

The internal inconsistency of this passage may suggest that Sherman was quite conscious that "the" common law upon which he based his bill did not in fact exist and that he was deliberately imposing a fictitious uniformity upon the precedent.

In his discussion of trusts Sherman identified consumer welfare as the policy of the common law in this area:

> "[W]hen [a combination] embraces the great body of all corporations engaged in a particular industry in all of the States of the Union, it tends to advance the price to the consumer of any article produced, it is a substantial monopoly injurious to the public, and, by the rule of both the common and the civil law, is null and void and the just subject of restraint by the courts . . . ."[91]

No senator challenged Sherman's representations of the common law, and two—Vest and Teller, the former a member of the Judiciary Committee—supported it. Speaking in support of Coke's amendment, Vest quoted the first section:

90 *Ibid.*

91 *Id.* at 2457.

> "That a trust is a combination of capital or skill by two or more persons, firms, or corporations for the purpose of creating or employing restrictions on trade, or limiting the production, increasing or reducing the price of merchandise or commodities, or preventing competition in the making, manufacture, sale or purchase of merchandise or commodities, or creating a monopoly in the manufacture, making, sale or purchase of any merchandise or commodity with intent to forestall the market value of any merchandise or commodity."

and stated, "There is a trust unlawful under the common law."[92] In fact, Vest went so far as to claim that a provision of Coke's bill which conditioned the application of its sanctions upon states declaring such trusts unlawful, was surplusage because of the uniformity of the common law on the topic in all states.[93]

Teller thought the states should attempt to reach the trusts with additional legislation, but he fully agreed with Sherman's statement of the law: "I understand that some of these trusts have been disturbed by the recent decisions of the courts of the country, which, as the Senator from Ohio [Sherman] showed the other day, have been all in one line, and I suppose no lawyer needs to have any argument made to him that these combinations and trusts are illegal without statute."[94]

> . . . THE TASK OF THE COURTS WOULD BE TO DISTINGUISH BETWEEN COMBINATIONS WHICH CREATE EFFICIENCY AND THOSE WHICH RESTRAIN TRADE.

There can hardly be any question that the discretion delegated to the courts by the Sherman Act was that of determining the consumer interest in particular cases and assessing legality accordingly. This is shown by Sherman's explicit statement that the task of the courts would be to distinguish between combinations which create efficiency and those which restrain trade. We have seen that by "restrain trade" Sherman meant "restrict output." The terms of the delegation are further shown by the policy of "the common law" which Sherman, Vest, and Teller, without contradiction, spelled out for the Senate.

92 *Id.* at 2603.

93 *Id.* at 2604. Vest's assertion had the incidental effect of equating Reagan's bill and the common law since Reagan and Coke defined "trusts" in language which was the same in all material respects.

94 *Id.* at 2560.

VII. The Absence of Expressed Values Other Than Consumer Welfare

Of those senators who supported the policy of Sherman's bill (as distinct from its constitutional footing or the remedies it provided), and they comprised the great majority of all who expressed views, not one suggested that the courts should in any case give weight to a value inconsistent with consumer welfare. It may be useful to examine some of the passages in the debates which have upon occasion been cited as expressive of conflicting views. A showing that these passages do not require, or in many cases even allow, such an interpretation should assist in establishing the intended exclusivity of the consumer welfare policy.

In the passage from the *Alcoa* opinion quoted first at the beginning of this paper, it will be recalled, Judge Hand attributed to Sherman "a desire to put an end to great aggregations of capital because of the helplessness of the individual before them." This helplessness was a noneconomic reason why "great industrial consolidations are inherently undesirable, regardless of their economic results." For this proposition Judge Hand relied upon two passages excerpted from Sherman's speeches. In the first, Sherman, speaking of trusts, said:

> "If the concentrated powers of this combination are intrusted to a single man, it is a kingly prerogative, inconsistent with our form of government, and should be subject to the strong resistance of the State and national authorities. If anything is wrong this is wrong. If we will not endure a king as a political power we should not endure a king over the production, transportation, and sale of any of the necessaries of life. If we would not submit to an emperor we should not submit to an autocrat of trade, with power to prevent competition and to fix the price of any commodity."[95]

It is at once apparent that Sherman's language not only fails to require Judge Hand's reading but refutes it. Sherman here analogizes the form of economic tyranny practiced by the trust to a political form, the "kingly prerogative." The latter is "inconsistent with our form of government," and so, by analogy, is the trust, the "autocrat of trade." If there were any doubt whatever about Sherman's meaning, it would be removed by the last sentence quoted. The thing which Sherman denounces is the "power to prevent competition and to fix the price of any commodity"—the power, in short, to injure consumers.

95 *Id.* at 2457.

The second passage quoted by Judge Hand came as part of a rhetorical crescendo in Sherman's opening speech urging adoption of his bill:

> "The popular mind is agitated with problems that may disturb social order, and among them all none is more threatening than the inequality of condition, of wealth, and opportunity that has grown within a single generation out of the concentration of capital into vast combinations to control production and trade and to break down competition. These combinations already defy or control powerful transportation corporations and reach State authorities. They reach out their Briarian arms to every part of our country. They are imported from abroad. Congress alone can deal with them, and if we are unwilling or unable there will soon be a trust for every production and a master to fix the price for every necessity of life."[96]

It is rather difficult to see what it is in this passage that might support the interpretation given it by Judge Hand. Either in or out of context, Sherman's words here are entirely consistent with his constant reference to the effect of the trusts upon consumers as the touchstone of illegality under his bill. Sherman does speak of inequalities of condition, wealth and opportunity, but it is abundantly clear that he does not suggest that the courts will or should use the law he proposes to create greater equality by dissolving large aggregations of capital regardless of the adverse impact this may have upon consumers by destroying efficiency. Sherman specifically complains only of those inequalities which are created by "the concentration of capital into vast combinations to control production and trade and to break down competition." He had explained already that these were to be forbidden because of their harmful impact upon consumers. If Sherman can be construed in this passage to welcome other forms of equality which would follow from the dissolution of monopolistic mergers, such results are clearly no more than a welcome by-product of a decision arrived at upon consumer welfare grounds. The same is clearly true of Sherman's remark that the combinations "reach State authorities." He was obviously not suggesting that, contrary to its explicit terms, the sanctions of his bill would be invoked upon proof that a trust had bribed or otherwise improperly influenced a state authority. The most that can be said of this passage is that Sherman took occasion to recount all of the sins of the trusts. To find in these words a mandate for a court to make a decision counter to the consumer welfare—in contradiction to everything else he had said on the topic—requires an effort beyond the merely heroic.

96 *Id.* at 2460.

Senator George frequently expressed concern over the plight of the small producer, and Judge Hand cited a page of one of his speeches which does contain some oratory that sounds as if it might support Judge Hand's thesis:

> "It is a sad thought to the philanthropist that the present system of production and exchange is having that tendency which is sure at some not very distant day to crush out all small men, all small capitalists, all small enterprises. This is being done now. We find everywhere over our land the wrecks of small independent enterprises thrown in our pathway. So now the American Congress and the American people are brought face to face with this sad, this great problem: Is production, is trade, to be taken away from the great mass of the people and concentrated in the hands of a few men . . . ."[97]

There is, in truth, a great deal of sympathy for small producers expressed in George's speeches. It seems abundantly clear, however, that George did not propose that the law's impact should ever be altered by that sympathy. He viewed the small producer interest and the consumer interest as complementary rather than conflicting and demanded action in the name of small producers only in situations where the same action would be required by the prevalent theory of consumer welfare. George agreed with Sherman's policy of protecting consumers,[98] and mentioned only two other situations as justifying, questions of constitutional power aside, the intervention of law: the imposition of lower prices upon small sellers by monopolistic combinations;[99] and the extraction of higher prices from small pro-

97 *Id*. at 2598. George was here recording his sentiments preparatory to announcing his inability to find any constitutional power in the Senate to deal with the problem. And, in fact, George thought that because of Congress' limited power the bill finally passed did not cover many cases. See note 82, *supra*. Thus, even if one thought that George did wish in the abstract to serve values that might conflict with consumer welfare, it would seem probable that he did not believe the Sherman Act would bear any such construction.

98 *Id*. at 1767-1768. George there complained that Sherman's double damage provision would not prove a sufficient encouragement to consumer lawsuits. Significantly, George read Sherman's bill as oriented entirely to consumer protection and his own bill employed the same tests as Sherman's. The difference between George's bill and Sherman's, aside from issues of constitutional power, lay in the remedies provided. This indicates that George was willing to have a law which was triggered only by injury to consumers but which, in such cases, had the further effect of protecting producers who were injured by the same cause. George's intent, therefore, could be carried out by applying only consumer-welfare criteria in the decision of cases.

99 Because he read Sherman's bill as requiring an intent to raise prices to consumers, George objected: "This leaves unpunished and perfectly lawful all those combinations which have proven so disastrous, that have for their object a decrease in the price to be given to the producer . . . ." *Id*. at 1767. Sherman met that objection by providing that a tendency toward a prohibited result would suffice to

ducers by monopolistic suppliers.[100] These are all cases which would call for the same legal intervention on a consumer-welfare rationale.[101] That George's concern for small producers was entirely complimentary to his concern for consumers is further shown by the fact that the bill he drafted employed precisely the same consumer-welfare criteria as Sherman's bill.[102] No man who proposed that the courts should favor producers over consumers in some cases would draft a law which made it illegal in every case to advance the cost of goods to consumers.[103] This reading also squares with George's participation as a member of the Judiciary Committee in the decision to permit monopoly gained by efficiency.

ANALYSIS DEMONSTRATES . . . THAT, WITH THE EXCEPTION OF THOSE FEW MEN WHO FAVORED A REASONABLE-PRICE TEST FOR CARTELS, IN NO CASE DID THE SPEAKER INTEND THAT COURTS IN DECIDING CASES SHOULD EVER PREFER THE PRESERVATION OF SMALL BUSINESS TO CONSUMER WELFARE.

There are scattered remarks by other legislators which might suggest to a casual reader that preservation of small business for its own sake was advocated. Analysis demonstrates, however, that, with the exception of those few men who favored a reasonable-price test for cartels, in no case did the speaker intend that courts in deciding cases should ever

footnote 99 cont'd
bring the law into play. *Id*. at 2461. George's case would thus be covered by the law since a combination to force purchase prices down can only operate by restricting purchases. This necessarily restricts the combination's output and injures consumers. Thus, once the issue of which injury was intended drops out, a consumer-welfare test also covers the wrong done to the small suppliers. George seems to have recognized this. Otherwise his objection to Sherman's bill on this point would have been broader than that the bill required an advance in prices to consumers to be intended. Later in the same speech George appeared to recognize explicitly this relationship between monopoly and monopsony: "[These trusts] operate with a double-edged sword. They increase beyond reason the cost of the necessaries of life and business and they decrease the cost of the raw material, the farm products of the country. They regulate prices at their will, depress the price of what they buy and increase the price of what they sell." *Id*. at 1768.

100 He gave the example of the cotton-bagging trust which injured "those small farmers, white and colored, who raise a few bales of cotton." *Id.* at 3147-3148. At times George referred to small producers who purchased from a trust as "consumers." *Id.* at 3149. Apparently he applied the term to anyone who purchased and did not resell in the same form. Ultimate consumers were merely part of a larger class. This does not affect the analysis, however, for any raising of prices to intermediate purchasers would raise the price to ultimate consumers as well.

101 See note 99 *supra.*

102 See p. 17 *supra.*

103 Those who did not wish to favor producers at the expense of consumers in some cases-such as in times of business depression-quite logically opposed Sherman's and Reagan's bills. See pp. 22-23 *supra.* George's bill, had it been seriously discussed, should have drawn precisely the same attack as Sherman's from these men. It is also significant that George, though he made almost every other conceivable attack upon Sherman's bill, never once joined those few who denounced its consumer orientation.

prefer the preservation of small business to consumer welfare.[104] Beyond this, it is impossible to find even colorable language suggesting most of the other broad

104 Shortly before the reference of S.1 to the Judiciary Committee, Edmunds said of the bill, then composed, in the part germane here, primarily of Sherman's draft plus Reagan's addition:

> I am in favor of the scheme [of the bill] in its fundamental desire and motive-most heartily in favor of it-directed to the breaking up of great monopolies which get hold of the whole of a particular business or production in the country and are enabled, therefore, to command everybody, laborer, consumer, producer and everybody else, as the sugar trust and the oil trust, and whatever. Although for the time being the sugar trust has perhaps reduced the price of sugar, and the oil trust certainly has reduced the price of oil immensely, that does not alter the wrong of the principle of any trust; and that in the brief definition of my friend from Texas [Reagan], is a phrase which covers every combination to get control of the life and the industry and the producing and the consuming classes of the country. I am in favor, most earnestly in favor, of doing everything that the Constitution of the United States has given Congress power to do to repress and break up and destroy forever the monopolies of that character, because in the long run, however seductive they may appear in lowering prices to the consumer for the time being, all human experience and all human philosophy have proved that they are destructive of public welfare and come to be tyrannies, grinding tyrannies, that have sometimes in other countries produced riots, just riots in the moral sense, and so on.

21 Cong. Rec. 2726.

Edmunds here identifies several categories of persons injured by the trusts, one of which is the category of consumers. He also states that the reduction of the prices of sugar and oil does not justify the trusts in those commodities. But evidence within and without this paragraph indicates that Edmunds was not advocating tests for legality not derived from a consumer-welfare rationale. Edmunds' language posits no conflict in the interests of these groups. Laborers, consumers, producers, and everybody else are injured by the trusts and abolition of trusts will benefit all of these groups. It would be entirely consistent with that position to construe the statute in accordance with consumer interests alone unless, in case of a conflict between the interests of the groups he named there was no reason to believe Edmunds would have preferred another group over consumers. Instead, there is reason to believe, however, that Edmunds assigned decisive weight to the consumer interest. His remark about the price of sugar and oil is consistent with this. The reduction in price is identified twice in the paragraph as a reduction only "for the time being." In "the long run" such monopolies will become "grinding tyrannies." This is consistent with the prevalent theory that monopolies are established by short-run low prices, which last only until rivals are ruined or join the trust, and are followed by low-run high prices. Edmunds' reference to the false seductiveness of the low prices of the trusts seems a recognition that the desirability of economic arrangements is properly judged by their effect upon consumers. Low prices are not seductive if your criterion is their effect upon rival producers.

This reading of Edmunds' remarks tends to be confirmed by his drafting of section 2 of the Sherman Act, see Thorelli, *op. cit. supra* note 8, at 212, his exchange with Kenna concerning the legality of monopoly by efficiency, see pp. 29-30 *supra*, and his acquiescence in Hoar's explanation of section 2, p. 30 *supra*. As already noted, monopoly by efficiency would appear to be disastrous to rival producers and as potentially tyrannous to labor as any other monopoly. It may be thought differ from other monopoly primarily in its net impact upon consumers. Thus, Edmunds much have considered the consumer interest decisive.

Mason, in the House of Representatives, however, may fairly be counted as preferring in some cases to protect small business at the expense of consumers: "Some say that the trusts have made products cheaper, have reduced prices; but if the price of oil, for instance, were reduced to one cent a barrel it would not right the wrong done to the people of this country by the 'trusts' which have destroyed legitimate competition and driven honest men from legitimate business enterprises." 21 Cong. Rec. 4100 (1890).

social or political purposes that have occasionally been suggested as relevant to the application of the Sherman Act.[105]

* * *

Having presented the evidence that leads me to conclude the Fifty-first Congress intended courts to apply a consumer-welfare policy exclusively, I turn to consider an objection. This is that the views of Senator Sherman and the discussion that turned on his and Reagan's proposals are irrelevant to the bill which the Judiciary Committee drafted and Congress enacted.[106]

The view that the debates which swirled around Sherman's draft of S.1 are largely irrelevant to the statute which ultimately emerged results from overestimating the severity of the break represented by the Judiciary Committee's redrafting of the bill. Walton Hamilton and Irene Till phrased this misunderstanding succinctly:

> "The great bother is that the bill which was arduously debated was never passed, and that the bill which was passed was never really discussed. . . . The

105 *E.g.,* Hoar, in asking that the Judiciary Committee draft be passed by the Senate without amendment, said, "The complaint which has come from all parts and all classes of the country of these great monopolies, which are becoming not only in some cases an actual injury to the comfort of ordinary life, but are a menace to republican institutions themselves, has induced Congress to take the matter up." 21 Cong. Rec. 3146 (1890). The reference to "the comfort of ordinary life" sounds like the consumer interest. The supposed threat of some trusts to "republican institutions" gives no reason to suppose that Hoar wanted courts to weigh such an imponderable standard in the decision of specific cases. He said nothing else of the sort, and his views on the commerce power, note 82 *supra,* and monopoly by efficiency, pp. 29-30 *supra*, preclude such an interpretation. Hoar was at most citing a by-product value of antitrust legislation. Bland in the House also denounced the effect of trusts on farmers but only in situations where the trust would also injure consumers. 21 Cong. Rec. 4099 (1890). The necessity to engage in analysis of such isolated remarks and shreds of casual rhetoric merely emphasizes how little there is in the legislative history to support theories that Congress intended courts to weigh social and political values other than consumer welfare in the application of the law.

106 I have not thought it worthwhile to consider in the text the often-heard statement that the Act must be construed in the light of the forces of Populism and agrarian discontent which are said to have provided much of the pressure for its passage. Not too much attention should be paid to such statements because they are essentially meaningless. Populism and the agrarian movements had not focussed on the general problem dealt with by the Sherman Act sufficiently to develop principles that a judge could apply predictably. This entire paper, moreover, is a refutation of the suggestion that Populist emotions, insofar as they might require a deviation from a consumer-welfare policy were enacted by Congress. Such emotions entered the debate on Ingalls' amendment, which was not enacted, and probably contributed to the sentiment in favor of exempting farm and labor organizations. But they left no traces elsewhere in the Sherman Act. Indeed, Sherman recognized the complaints of farmers and workingmen but said, "They can not see the cause or source of this evil . . . ." 21 Cong. Rec. 2569 (1890). See also, Edmunds, *id.* at 2728, 3148; Hoar, *id.* at 2568. The men who shaped the Sherman Act undoubtedly felt the pressure of popular discontent with the "trusts," but they chose their own remedy.

[Judiciary] committee turned a deaf ear to all that the Senate had said and done and went its own way. Intent, therefore, forsakes the Congressional Record for the capacious recesses of that flexible corpus called the common law."[107]

The authors also state that the law "is to this day strangely enough called the Sherman Act—for no better reason, according to its author [Hoar], than that Senator Sherman had nothing to do with it whatever."[108] These assertions are provocative, to say the least, and it is unfortunate that Hamilton and Till do not indicate the evidence upon which their statements rest.[109]

My own study of the Congressional Record leads me to conclude that the policy of the bill so "arduously debated" was carried forward into the Judiciary Committee's draft and enacted. The popular name of the statute correctly attributes paternity to Sherman. There is no reason to doubt this other than the fact that the Judiciary Committee recast S.1 in common law terms. Common sense alone, moreover, makes the Hamilton-Till thesis dubious. The shortness of the Senate debate over the Judiciary Committee's redraft certainly suggests that the Senate thought it knew what the draft meant, and that can be explained only on the theory that the proceeding discussions of Sherman's policy were fully applicable to the new draft. In fact, when Hoar brought the redraft in he told the Senate, "I shall not undertake to explain the bill, which is well understood."[110] There are other good reasons to believe that the Senate thought the Judiciary Committee draft represented the basic polices espoused by Sherman. He was by far the most articulate and thorough speaker on the question of what goals antitrust should serve. Those who spoke overwhelmingly agreed with his position on this issue. Disagreement was largely confined to questions of remedies and the constitutional reach of Sherman's measure. The reference to the Judiciary Committee was finally made, after having been voted down twice, because of concern with those matters as well as the meaning and constitutionality of the various additions, such as the Ingalls amendment, which had been

107 Hamilton and Till, Antitrust in Action 11 (TNEC Monograph No. 16, 1940).

108 *Id*. at 10. The authors cite Hoar's autobiography for his claim. Thorelli and others have since identified Edmunds as the principle draftsman of the statute in its final form. Thorelli, *op. cit. supra* note 8, at 210-214.

109 They state only that their materials came very largely from the Senate and House bills and from the debates in the 51st Congress as reported in the Congressional Record, but "to equip each sentence, almost each phrase, with its particular citation would be as cumbersome as it is unnecessary." Hamilton and Till, *op. cit. supra* note 107, at 5.

110 21 Cong. Rec. 3145 (1890).

made to S.1.[111] The one major issue which arose in connection with the Judiciary Committee draft, the issue of monopoly due to efficiency, was, moreover, explained by Hoar and Edmunds and resolved in a manner consistent with Sherman's consumer-welfare rationale.

Even more clear-cut evidence is supplied by the role of "the common law" in the Senate's deliberations. Prior to the Judiciary Committee reference Sherman's bill and policy were firmly and repeatedly identified with the common law. Sherman gave the Senate an extended discussion of common law cases and principles which he said his bill would enact for federal enforcement. He repeatedly used common law terminology, "restraint of trade," as interchangeable with his bill's reference to prevention of full and free competition and advancement of costs to consumers.[112] Even the title of his bill made the point: "A bill to declare unlawful trusts and combinations in restraint of trade and production."[113] That title not only identified the consumer-welfare tests of the bill with the common law of restraints of trade, but, by adding the term "restraint . . . of production" suggested that the evil was restriction of output. We have seen that Sherman made the same point in defending his bill as a proper exercise of the commerce power and in identifying the mechanism by which trust's advanced costs to consumers. No senator challenged Sherman's version of the common law or his assertion that his bill merely enacted it. Senators Vest and Teller explicitly agreed with Sherman.

That title not only identified the consumer-welfare tests of the bill with the common law of restraints of trade, but, by adding the term "restraint . . . of production" suggested that the evil was restriction of output.

When the Judiciary Committee, which had not been asked to alter or amend Sherman's policy in any way, reported back a redraft that made the test of illegality the "restraint of trade or commerce" members of the Senate had every reason to think that this use of the common law phrase carried Sherman's policy views with it. This is particularly true because Vest, who had agreed with Sherman on this point, was a member of the committee. In reporting the redraft, moreover, Hoar three times identified it with the common law. He said the committee had "affirmed the old doctrine of the common law in regard to all inter-

111 See, for example, 21. Cong. Rec. 2597-2611, 2655, 2656-2657, 2659-2660, 2731 (1890). The monstrosity which S.1 had become by amendment and the confusion which surrounded it prior to its commitment to the Judiciary Committee can be seen by comparing the reprint of March 25, 1890, Bills and Debates, 217, with that of March 26, *id*. at 227, and then reading 21 Cong. Rec. 2723-2726 (1890) up to Edmunds' speech. The confusion which attended Ingalls' amendment is reflected in the discussion, *id.* at 2646-2662.

112 *E.g., id.* at 2456, 2457, 2462.

113 Bills and Debates 69.

state and international commercial transactions."[114] In replying to Kenna on monopoly by efficiency he said the offense of monopolizing prohibited by section 2 of the bill was defined by the common law,[115] and, in an argument very like that of Sherman's, he said:

> "The common law in the States of the Union of course extends over citizens and subjects over which the State itself has jurisdiction. Now we are dealing with an offense against interstate and international commerce, which the State can not regulate by penal enactment, and we find the United States without any common law. The great thing that this bill does, except affording a remedy, is to extend the common law principles, which protected fair competition in trade in old times in England, to international and interstate commerce in the Untied States."[116]

Edmunds, too, said the committee had decided to "make its [the bill's] definition out of terms that were well known to the law already" and leave it to the courts "in the first instance to say how far they could carry it or its definitions as applicable to each particular case as it might arise."[117] He said the bill "is clear in its terms, is definite in its definitions, and is broad in its comprehension,"[118] statements he could hardly have made had the debates prior to the Judiciary Committee reference not been thought to give content to the common law terminology of the final bill. Morgan expressed the same thought and appeared also to recognize that "the common law" was being made to say new things: "[W]e use common-law terms here and common-law definitions in order to define an offense which is in itself comparatively new . . . ."[119]

The Senate was thus told what "the common law" was and then repeatedly assured that both Sherman's bill and the Judiciary Committee's redraft were enactments of that law. According to a well-known axiom, Sherman's policy and

114 21 Cong. Rec. 3146 (1890).

115 *Id.* at 3152.

116 *Ibid*.

117 *Id.* at 3148.

118 *Ibid.*

119 *Id.* at 3149.

the Judiciary Committee's policy, being equal to the same thing, are equal to each other.

* * *

The contract we call "legislative intent" must be used with care. If for no other reason than its inherent artificially, "legislative intent" cannot properly be used to settle all questions about the bounds of judicial discretion. I offer this paper, therefore, less to demonstrate that Sherman Act issues are only those relevant to consumer welfare—through such weight as "legislative intent" may have surely pulled in that direction—than to rebut contrary claims which purport to rest upon a discernible congressional intention. If values other than consumer welfare are to be made legitimate criteria for Sherman Act litigation, the legitimation will have to proceed from some base other than the "purpose" of the Fifty-first Congress.

It is difficult to resist the conclusion that the most faithful judicial reflection of Senator Sherman's and his colleagues' policy intentions was the rule of reason enunciated by Chief Justice White in the 1911 *Standard Oil* and *American Tobacco* opinions. There was in White's opinions as in Sherman's speeches the idea that the statute was concerned exclusively with consumer welfare and that this meant the law must discourage restriction of output without hampering efficiency.[120] White appears also to have incorporated into his rule of reason those major rules of law which Sherman envisaged as implied by a consumer-welfare policy.[121] The rules implied by the policy are alterable as economic analysis progresses, however. White clearly foresaw this and incorporated that principle of change into his rule of reason.[122]

Courts charged by Congress with the maximization of consumer welfare are free to revise not only prior judge-made rules but, it would seem, rules contemplated by Congress. The Sherman Act defines the class of situations to which it may be applied, but it does not freeze into statutory commands the rules of legality about predation, mergers, and so forth, that many congressmen contemplated. Sherman and others clearly believed that they were legislating a policy and delegating to the courts the elaboration of subsidiary rules. Nothing in the legislative history or in the language of the statute suggests that courts are required to hold any specific type of agreement or behavior unlawful regardless of its prob-

120 See Bork, The Rule of Reason and the Per Se Concept: Price Fixing and Market Division, 74 Yale L.J. 775, 801-805, 829-832 (1965), and 75 Yale L.J. 373, 375-377, 375 n.2 (1966).

121 That is, rules against cartel agreements, monopolistic mergers, and predatory practices. See 74 Yale L.J. at 801-805.

122 *Id.* at 802, 805.

able impact upon consumers. In terms of "law," therefore the Sherman Act tells judges very little. A judge who feels compelled to a particular result regardless of the teachings of economic theory deceives himself and abdicates his delegated responsibility. That responsibility is nothing less than the awesome task of continually creating and recreating the Sherman Act out of his understanding of economics and his conception of the requirements of the judicial process.

Report by the EAGCP

"An economic approach to Article 82"

July 2005

Jordi Gual

IESE Business School and "la Caixa", Barcelona

Martin Hellwig

Max Planck Institute for Research on Collective Goods, Bonn

Anne Perrot

University Paris I and Conseil de la Concurrence, Paris

Michele Polo

Bocconi University, Milan

Patrick Rey (Coordinator)

University of Toulouse

Klaus Schmidt

University of Munich

Rune Stenbacka

Swedish School of Economics, Helsinki and RUESG, University of Helsinki

Executive Summary

This report argues in favour of an economics-based approach to Article 82, in a way similar to the reform of Article 81 and merger control. In particular, we support an effects-based rather than a form-based approach to competition policy. Such an approach focuses on the presence of anti-competitive effects that harm consumers, and is based on the examination of each specific case, based on sound economics and grounded on facts.

Why do we need an economic approach?

An economic approach to Article 82 focuses on improved consumer welfare. In so doing, avoids confusing the protection of competition with the protection of competitors and it stresses that the ultimate yardstick of competition policy is in the satisfaction of consumer needs. Competition is a process that forces firms to be responsive to consumers' needs with respect to price, quality, variety, etc.; over time it also acts as a selection mechanism, with more efficient firms replacing less efficient ones. Competition is therefore a key element in the promotion of a faster growing, consumer-oriented and more competitive European economy.

An economics-based approach requires a careful examination of how competition works in each particular market in order to evaluate how specific company strategies affect consumer welfare. Indeed, an economic approach achieves two complementary goals. First, it ensures that anti-competitive behaviour does not outwit legal provisions. By focusing on the effects of company actions rather than on the form that these actions may take, an economics-based approach makes it more difficult for companies to circumvent competition policy constraints by way of attempting to achieve the same end results through the use of different commercial practices. At the same time, this approach provides a more consistent treatment of practices, since any specific practice is assessed in terms of its outcome and two practices leading to the same result will therefore be subject to a comparable treatment.

Second, the economics-based approach guarantees that the statutory provisions do not unduly thwart pro-competitive strategies. An effects-based analysis takes fully into consideration the fact that many business practices may have different effects in different circumstances: distorting competition in some cases and promoting efficiencies and innovation in others. A competition policy approach that directly confronts this duality will ensure that consumers are protected (through the prevention of behaviour that harms them) while promoting overall increased productivity and growth (since firms will not be discouraged in their search for efficiency).

What are the implications of an economic approach?

An economics-based approach to the application of article 82 implies that the assessment of each specific case will not be undertaken on the basis of the form that a particular business practice takes (for example, exclusive dealing, tying, etc.) but rather will be based on the assessment of the anti-competitive effects generated by business behaviour. This implies that competition authorities will need to identify a competitive harm, and assess the extent to which such a negative effect on consumers is potentially outweighed by efficiency gains. The identification of competitive harm requires spelling out a consistent business behaviour based on sound economics and supported by facts and empirical evidence. Similarly, efficiencies – and how they are passed on to consumers– should be properly justified on the basis of economic analysis and grounded on the facts of each case.

An economics-based approach will naturally lend itself to a "rule of reason" approach to competition policy, since careful consideration of the specifics of each case is needed, and this is likely to be especially difficult under "per se" rules. At the same time, we should not fall into the trap of active intervention and fine-tuning; whenever possible, competition is to be preferred to detailed regulation as the best mechanism to avoid inefficiencies and foster productivity and growth; this calls for a "non-dirigiste" approach to competition policy that focuses in most cases on entry barriers; in the context of Article 82, it is then natural to focus on competitive harm that arises from exclusionary strategies –possible exceptions concern some natural monopoly industries which may require ongoing supervision of access prices and conditions by regulatory agencies. Without trying to be exhaustive, the report discusses well-identified exclusionary strategies and the practices that they involve.

What are the consequences for procedure and the predictability of competition policy?

In terms of procedure, the economic approach implies that there is no need to establish a preliminary and separate assessment of dominance. Rather, the emphasis is on the establishment of a verifiable and consistent account of significant competitive harm, since such an anti-competitive effect is what really matters and is already proof of dominance. In an effects-based approach, the focus is on the use of well-established economic analysis. Such a conceptual framework provides a benchmark for the detailed assessment of the key ingredients that have to be present in a case, whether one tries to check the presence of significant competitive harm, or the achievement of relevant economic efficiencies.

This approach has also natural implications in terms of the burden of proof in specific cases. Competition authorities have to show the presence of significant anti-competitive harm, while the dominant firm should bear the burden of establishing credible efficiency arguments.

Requiring consistent economic arguments, grounded on established facts, may be perceived as constraining somewhat the competition authority' leeway. It is however necessary to ensure a consistent treatment of the various practices that can serve the same anticompetitive effect. It also contributes to enhance the predictability and, consequently, the effectiveness of competition policy enforcement. Indeed, under a form-based approach predictability need not be higher (as the Michelin saga shows), and sometimes the predictability inherent to "ex-ante" prohibitions may in fact be a straightjacket for business, preventing innovation and economic growth.

Overall, we believe that the economics-based approach presented in this report is a step in the right direction: focused on consumers' needs and the promotion of economic growth, and consistent with the reforms of article 81 and merger control. It is not a question of having more or less intervention, but of more effective intervention. The goal is to focus on the important competitive harms, while preserving and encouraging efficiency. The economic approach to article 82 provides a flexible framework that fosters increased productivity and growth to the benefits of consumers.

Chapter I: General principles

Section 1: Effects-based versus form-based approach

European competition policy has recently been reformed and is now following a more economics-based approach in the implementation of Article 81 on anticompetitive agreements and in merger control; we will argue here in favour of a similar move for the enforcement of Article 82 on abuses of dominance. In particular, we will argue that the competition authority should adopt an effects-based rather than a form-based approach to competition policy. In this chapter we first review the net benefits of such a move, and then make a few remarks on how to implement it.[1]

The discussion and management of Article 82 cases are often organized by categories of conduct, such as predatory pricing, discrimination, fidelity rebates or tying. However, such a form-based approach is problematic. In many instances alternative practices can serve the same purpose. For example, predatory pricing can take the form of selective rebates, targeted at the rival's prospective customers. Alternatively, the predator can engage in explicit discrimination and charge more attractive prices or, more generally, offer better conditions to these customers. Other instruments in the predator's toolbox include implicit discrimination (e.g. in the form of fidelity or quantitative rebates that are formally available to all, but in fact tailored to the specific needs of the targeted customers) and mixed bundling or tying, when these customers are particularly interested in the bundle in question. To take another example, a firm that controls a key input may distort competition in a downstream market by refusing to deal with independent downstream firms; alternatively, it can engage in exclusive dealing arrangements or engage in explicit or implicit price discrimination such as mentioned above; yet other instruments include specific (in-)compatibility choices, physical or commercial tying, and so forth.

[1] Vickers, J. (2004), Abuse of Market Power, *The Economic Journal*, 115 (June): F244–F261, offers a detailed discussion of enforcement issues.

A more consistent approach would start out from the effects of anticompetitive conduct, such as exclusion of competitors in the same market or in a horizontally or vertically related market one, and consider the competitive harm that is inflicted on consumers. Adopting such an effects-based approach would ensure that these various practices are treated consistently when they are adopted for the same purpose. In contrast, a form-based approach creates the risk that they will be treated inconsistently, with some practices possibly enjoying a relatively more lenient attitude (e.g., because of different standards). Arbitraging among these different treatments may facilitate exclusion, or induce the dominant firm to adopt alternative exclusionary methods, which may well inflict a higher cost on consumers.

For example, in the context of predation, tight rules against predatory pricing may lead the predator to offer better terms in other dimensions. It may be less likely that these better terms are passed on to final consumers. Final consumers then will no longer benefit from low prices in the short-run, and yet this alternative strategy may still have similar exclusionary effects in the long-run. Similarly, in the context of vertical foreclosure, banning discrimination may lead the bottleneck owner to refuse to deal with any independent downstream competitor – which may well be interpreted as a more extreme form of discrimination; by this decision, however, the dominant firm gives up somewhat on product diversity. This reduces its profitability but also hurts consumers.

While alternative practices can serve the same purpose in given circumstances, the same practice can also have either pro- or anticompetitive effects, depending on the circumstances. To take a simple example, low prices are at the heart of a desirable competitive process. And in some cases (introductory pricing, economies of scale and scope, learning-by-doing, network effects), even prices that are below cost for some period constitute "normal" competitive prices. But it is also true that in specific circumstances, low prices (even above cost) can have an exclusionary effect and harm consumers in the medium- to long-run. There again, focussing on effects, as opposed to form, is key to an effective competition policy.

Whether we consider an effects-based or a form-based approach, in the enforcement of competition policy, one can opt for *per se* rules, i.e., an ex-ante description of what is banned, or for a rule of reason, i.e. an ex-post overall evaluation of the different consequences. In an effects-based approach a *per se* rule for "financial predation", for example, would be "whenever you have a financially strong incumbent/financially weak entrant, the incumbent cannot 'invest in losses' beyond a certain threshold (to be defined with reference to possible practices),"; in contrast, a form-based *per se* approach to "predatory pricing" would prescribe that "the incumbent cannot lower its price below a certain threshold." That is, even when implemented as a *per se* rule, an effects-based approach includes more than a single practice (low pricing, high advertising, and so forth) and requires a richer description of the circumstances (financial conditions of the incumbent and the competitor) than a form-based approach. In both cases, however, the competition authority must balance the likelihood of false positives (condemning a pro-competitive practice in a particular case) and false negatives (allowing a dominant firm to abuse its market power in other cases), as well as the likely magnitudes of the costs for competition of both types of errors. Thus, if, for example, a *per se* approach were to be adopted, the *per se* rule should be a systematic ban only if the expected cost of false negatives is perceived to dominate. The economic approach highlighted below suggests however that in general both types of errors are likely, and that they can be much more accurately assessed when taking into account the specific circumstances of a case. This therefore provides a strong argument in favour of a rule of reason. Still, the standard and the process adopted should be designed to balance these two types of errors; for example, if the cost of false negatives is expected to be higher, then the balance should be tilted towards plaintiffs and against dominant firms. In a form-based approach, however, this balancing would be done for each practice. This could indeed lead to an inconsistent treatment of alternative practices, even though they may be used for the same purpose. In contrast, in an effects-based approach the balance would be made according to the type of anticompetitive harm that is at stake. This would not only ensure that treatment is more consistent, but it would also focus the attention and the scarce resources of competition authorities on those cases where competitive harm is likely to be important.

Section 2: Objectives

a) Consumers matter

In the preceding account of competition policy, we have asked that the competition authority start by identifying the competitive harm that is involved in the case under review. This begs the question of what is to be meant by "competitive harm". This question in turn is linked to the question of what are the objectives of competition policy. In one tradition the objective of competition policy is defined as the protection of competition.[2] This formula as such is not very helpful because it raises the further question of the standards by which the competition authority is to assess a given type of conduct in practice. If a particular type of conduct permits a company to succeed and to displace its competitors in the market, by what standard should the competition authority assess whether the conduct in question is detrimental to "competition" or whether the conduct in question is legitimate and its prohibition by the authority would be detrimental to "competition"? Ultimately, the assessment of competitive harm must be based on an assessment of how competition in the particular market works and what the practice in question means for market participants.

The standard for assessing whether a given practice is detrimental to "competition" or whether it is a legitimate tool of "competition" should be derived from the effects of the practice on consumers. If we think of "competition" as a regime in which the different suppliers contend to sell their products to participants on the other side of the market, then the benefits reaped by the other side of the market will themselves provide a measure of how well "competition" works. For final-products markets, this observation leads directly to a consumer welfare standard. For primary- or intermediate-products markets, a consumer welfare standard is obtained by adding the observation that the vertical organization of industry itself is a subject of "competition" the ultimate beneficiaries of which are the final consumers. In either case, competition forces the supply side of the economy to be responsive to consumers' needs with respect to price, quality, variety, etc.; business strategies that respond to these needs and raise consumer welfare are likely to be legitimate competitive strategies. The observation of such a strategy in the market provides prima facie evidence of the importance of competition. In contrast, a lowering of consumer welfare provides evidence of competitive harm.

If the assessment of competitive harm and the protection of "competition" are assessed with reference to consumer welfare, it is incumbent upon the competition authority in each case to examine the actual working of competition in the particular market without prejudice and to

---

[2] Thus, in its XXIX Report on Competition Policy (2000, p. 6), the European Commission writes "The first objective of competition policy is the maintenance of competitive markets."

explain the harm for consumers from the practice in question. Without the discipline provided by this routine, the authority may be tempted to identify the "protection of competition" with the preservation of a particular market structure, e.g., one that involves actual competition by a given company. Its policy intervention may then merely have the effect of protecting the other companies in the market from competition. This would enable them to maintain their presence in the market even though their offerings do not provide consumers with the best choices in terms of prices, quality, or variety.[3]

In some cases, concerns for the protection of competitors from certain forms of inappropriate behaviour may be appropriate. However, this is certainly not true for all cases; moreover, competitors themselves should not be protected from competition by the authority's intervention. In each case, the competition authority must assess these matters without prejudice for any particular structure. A consumer welfare standard in the context of an effects-based approach provides a suitable criterion for distinction. Referring to this standard is all the more important because, in the actual proceedings on a given case, competitors are usually much better organized than consumers. The competition authority receives more complaints and more material from competitors, so the procedure tends to be biased towards the protection of competitors. Developing a routine for assessing consumer welfare effects provides a counterweight to this bias.

b) Tradeoffs, Proxies, and the Need for a Non-Dirigiste Approach

The use of the single term "consumer welfare" conceals the fact that we are really talking about a multi-facetted concern. Many issues concern multiple markets, with consumer welfare effects going in different directions in the different markets. In many cases, it is also necessary to think about consumer welfare in the future, as well as consumer welfare today.

In such cases, the competition authority needs to take a comprehensive view, taking account of the different effects of the practices under investigation and of policy interventions on consumer welfare. In particular, it must allow for long-term, as well as short-term effects, and

[3] Thus, in the Lufthansa-Germania predatory pricing case, in November 2001, the Bundeskartellamt required Lufthansa to keep its price on the Düsseldorf-Berlin route at least 35 Euro above the price of 99 Euro at which the competing carrier Germania claimed to be covering its costs. The difference was deemed to reflect the costs of services provided by Lufthansa, but not by Germania. The relation of the base price of 99 Euro to Lufthansa's costs was not considered (though Lufthansa's claims in other cases suggest that the costs were higher).

for repercussions on neighbouring markets, as well as the markets under direct consideration. A neglect of longer-term effects or indirect effects is not justified by the fact that these effects are more uncertain and more difficult to assess.

In considering consumer welfare effects in multiple, present, and future markets, one usually faces tradeoffs. In some cases, the markets themselves provide information about the appropriate weights to be attached to the different effects, e.g. discount factors for weighing future as opposed to present effects. In many cases though, the tradeoffs involve an element of redistribution between groups of consumers. For such redistribution, market data – or indeed any other data – do not provide "objective" guidance. When faced with such tradeoffs, the competition authority must exercise its judgement, which necessarily involves a certain element of subjectivity. However, acknowledging that there is room for a certain element of subjectivity in taking choices concerning tradeoffs does not absolve the competition authority from the requirement to be clear about the tradeoffs themselves and to indicate precisely what consumer welfare effects are relevant to its decision.

In focussing on consumer welfare, one must not fall into the trap of seeing competition policy as a tool of active policy intervention designed to correct the inefficiencies associated with monopolies and oligopolies so as to maximize some measure of welfare. Competition policy is based on the principle that competition itself is the best mechanism for avoiding inefficiencies, so the competition authority should not try to let its own intervention replace the role of competition in the market place. The powers given to the competition authority are, with very few exceptions, powers to prohibit certain behaviours and certain developments, not powers to actively determine where the market participants should be going. The authority can ban certain agreements, certain practices and certain mergers, but it should not tell the markets participants what they should do instead.

As an illustration of these considerations, consider the problem of monopoly pricing. One response to the problem might be for the competition authority to intervene, citing excessive pricing by a monopolist as an infraction of the abuse-of-dominance prohibition in Article 82 of the Treaty. Another response might be to leave the matter alone, hoping that the profits that the monopolist earns will spur innovation or imitation and entry into the market, so that, eventually, the problem will be solved by competition.

The choice between these two alternative responses to the problem of monopoly pricing involves a choice among competition policy regimes, as well as an intertemporal tradeoff. If it was just a question of short-run versus long-run effects, one might be tempted to put the immediate gain of today's consumers above everything else. However, a policy intervention on such grounds requires the competition authority to actually determine what price it considers appropriate, as well as how it should evolve over time; for this it is not really qualified. Moreover, such a policy intervention drastically reduces, and may even forego the chance to protect consumers in the future by competition rather than policy intervention. A regime in which consumer protection from monopoly abuses is based on competition is greatly to be preferred to one in which consumer protection is due to political or administrative control of prices. In most circumstances therefore, the competition authority ought to refrain from intervening against monopolistic pricing and instead see to it that there is room for competition to open up.

Exceptions to this recommendation tend to involve monopolies that own essential facilities such as transmission and distribution grids in electricity, whose reduplication is technically infeasible or economically undesirable. Policy intervention to control the use of monopoly power derived from such facilities can be desirable, particularly as a tool for enhancing competition in activities such as the generation and sale of electricity, which are not themselves "natural monopolies", but require access to the essential facilities. However, in designing the appropriate rules for access provision and for the pricing of such facilities, one must provide for ongoing supervision. This falls outside the scope of traditional competition policy and is best left to specialized regulatory authorities. By no means should the justifiability of a dirigiste approach for some essential facilities become a paradigm for competition policy itself.

In assessing the implications of alternative policies for the future, one difficulty is that their effects on future market outcomes are difficult to predict. Trying to foresee the different possibilities is sometimes quite hopeless, especially if one takes into account that the genius of competitive markets lies precisely in developing possibilities that no one has thought about before. Given this difficulty, it is sometimes necessary to forego an explicit computation of consumer welfare in future markets and to rely on a proxy instead. Such a proxy may usefully refer to aspects of market structure.

Thus in the preceding discussion of the monopoly problem, a structure in which the policy authority foregoes any attempt to control monopoly pricing, but other suppliers are free to enter the market was deemed superior in most cases to a structure involving policy control of monopoly pricing and reduced entry incentives. This judgement is based on broad patterns of experience concerning the implications of the two alternatives in a variety of markets rather than any attempt at making precise predictions for the market under consideration. As such, it conforms to the anti-dirigiste approach to competition policy.

In the monopoly example, the use of structural aspects of the market, namely the freedom of entry and the prospect of future competition, as a proxy for the explicit assessment of consumer welfare effects in future markets serves to caution the competition authority about the dangers of intervention. In other examples, e.g., in cases involving market foreclosure, the use of such a proxy can also work in the opposite direction and call for a policy intervention designed to prohibit exclusionary practices and to keep markets open for competition. The structural aspects of markets that need to be taken into account encompass traditional notions of market structure, but extend far beyond them, including, in particular, the potential for entry by new competitors. In each case, the competition authority must give a clear account of (i) the precise material justification for treating the structural effects in question as a proxy for future consumer welfare effects and (ii) the precise way the business practice under investigation affects the scope for current and future competition. Particular attention must be paid to exclusionary effects restricting the scope for new entry. Competitive harms from exclusionary effects are discussed in detail below. However, there should be no prior presumption that the current market structure and the current competitors are the guarantors of competition, which will enhance consumer welfare in markets in the future.

Section 3: Procedure

Moving from a form-based to an effects-based approach has important implications for procedure. Whereas under a form-based approach, it is enough to verify (i) that a firm is dominant and (ii) that a certain form of behaviour is practiced, an effects-based approach requires the verification of competitive harm.

In the first place, in deciding to bring a case, the competition authority should therefore focus on identifying the competitive harm of concern. To do so, the authority must analyse the practice in question to see whether there is a consistent and verifiable economic account of significant competitive harm. The account should be both based on sound economic analysis and grounded on facts. In particular, since many practices can have pro- as well as anticompetitive effects, merely alluding to the possibility of a story is not sufficient. The required ingredients of the story must therefore be properly spelled out and shown to be present. At the same time, the authority must check to see whether the practice in question cannot also be justified as a legitimate mode of competitive behaviour. If several interpretations are possible, the authority must investigate whether the data permit a distinction as to which of the different interpretations apply.

In asking for an account that is based on sound economic analysis and grounded on facts, we are *not* asking for an account that embeds the case under review into a single, encompassing formal model. Formal models are designed to verify the consistency of arguments about one particular effect and to gain an idea about the empirical data one needs in order to assess the relevance this effect in a given case. In any given case in practice, however, one may have to examine several effects at once; in this situation, an encompassing formal analysis may not be feasible. However, for each particular effect that is considered, the arguments that are made should be grounded in formal analysis. At this level the analysis should rely on models as tools to assess the validity of the argument – in its relation to the facts, as well as internal consistency, and consistency with the other arguments that are given. Where empirical information points to effects that have not yet been studied in the literature, it may be necessary to develop a model from scratch, relying on standard methods. However, the less a given effect has been scrutinized in theoretical and empirical research, the more cautious the authority should be in relying on it for the account which it gives.

Requiring this first step may be perceived as constraining somewhat the competition authority's leeway. It is however necessary to ensure the consistency of the treatment of the various practices that can serve the same anticompetitive effect. It also contributes to enhancing the predictability and, consequently, the effectiveness of competition policy enforcement.

In contrast to a form-based approach, an effects-based approach needs to put less weight on a separate verification of dominance, except possibly for a de minimis consideration. If an effects-based approach yields a consistent and verifiable account of significant competitive harm, that in itself is evidence of dominance. Traditional modes of establishing "dominance" by recourse to information about market structure are merely proxies for a determination of "dominance" in any substantive sense, i.e., the ability to exert power and impose abusive behaviour on other market participants. If an effects-based approach provides evidence of an abuse which is only possible if the firm has a position of dominance, then no further separate demonstration of dominance should be needed – if no separate demonstration of dominance is provided, one may however require the abuse to be clearly established, with a high standard of proof

Traditional considerations about the presence or absence of dominance do not therefore become moot. They merely become part of the procedure for establishing competitive harm by the practice under investigation. Thus, in a predation case, any account of the possibility of recouping current losses through future gains will have to involve some verification of the firm's prospects for imposing and maintaining higher prices once the presumed predation has been successful.

In proposing to reduce the role of separate assessments of dominance and to integrate the substantive assessment of dominance with the procedure for establishing competitive harm itself, we depart from the tradition of case law concerning Art. 82 of the Treaty, but *not*, we believe, from the legal norm itself. Art. 82 of the Treaty is concerned not just with dominance as such, but with *abuses* of dominance. The case law tradition of having separate assessments of dominance and of abusiveness of behaviour simplifies procedures, but this simplification involves a loss of precision in the implementation of the legal norm. The structural indicators which traditionally serve as proxies for "dominance" provide an appropriate measure of power in some markets, but not in others. In a market where these indicators do not properly measure the firm's ability to impose abusive behaviour on others, the competition authority's intervention under traditional modes of procedure is likely to be inappropriate, too harsh in some cases and too lenient in others. Given that the Treaty itself does not provide a separate definition of dominance, let alone call for any of the traditionally used indicators as such, it seems more appropriate to have the implementation of the Treaty itself focus on the abuses and to treat the assessment of dominance in this context.

Our proposed effect-based approach also allows us to capture in a balanced and meaningful way the notion of special responsibility of a dominant firm. The reference to such responsibility is often intended to prohibit some practices when exerted by a dominant firm, while considering them lawful if practiced by smaller competitors. Once we focus on the exclusionary effects of market practices, the notion of special responsibility naturally emerges from the analysis, in that certain practices are to be prohibited when they determine exclusionary effects, while they are lawful as long as no competitive harm is involved. Since in this analysis we do not need to assess the existence of dominance separately, the special responsibility implicitly applies to any conduct and firm that (is able to) interfere and distort the competitive process of entry into the market.

Moving from a form-based to an effects-based approach will pose a challenge for court proceedings. A natural process would consist of asking the competition authority to first identify a consistent story of competitive harm, identifying the economic theory or theories on which the story is based, as well as the facts which support the theory as opposed to competing theories. Next, the firm should have the opportunity to present its defense, presumably to provide a counter-story indicating that the practice in question is not anti-competitive, but is in fact a legitimate, perhaps even pro-competitive business practice. In the end, it will be up to the court to determine which story it considers to be the most plausible.

Given the creativity of lawyers and economists in coming up with stories, the outcome of such proceedings can be very sensitive to how the burden of proof is allocated between the two parties. In line with the procedure sketched out above, the general rule should be that the antitrust authority bears the burden of proof for identifying and establishing anticompetitive effects. Two additional principles seem advisable: First, in the absence of additional evidence to the contrary, an argument based on established economic theory and supported by facts that according to the theory, are material to the assessment of the practice in question should be deemed more credible than a counterargument that does not have such a basis. For example, in a case involving financial predation, it should be enough for the competition authority to establish the circumstances which, according to existing economic theory, can make financial predation viable without having to provide a more detailed account of the possibility of recoupment.[4] Second, if the story of competitive harm that is brought forth by the competition authority fulfils the criteria listed above and the validity of the counter-story brought by the firm hinges on data in the domain of the firm, then it should be incumbent upon the firm to provide these data.

Both of these principles require a certain degree of flexibility in the handling of proof requirements. However, this cannot be avoided if the effects-based approach is to be practical. Whereas a form-based approach hinges on data that the competition authority should in principle be able to provide, an effects-based approach also requires interpretations of data where discrimination is more difficult. Taking existing economic theory as a standard of reference removes some arbitrariness from such an interpretation.

The greater flexibility of an effects-based approach need not reduce the predictability of competition policy. To be sure, within a given tradition, a form-based approach may provide the competition authority as well as market participants with some guidance as to what forms of behaviour are acceptable and what forms are not. However, short of stipulating an exhaustive list of acceptable forms of behaviour, which would probably constitute an uncomfortable "straight jacket" that would impede economic progress and development, the case law will only determine a list of practices that are not acceptable. By nature, that list cannot be exhaustive: firms that want to purse a given purpose (be it pro- or anti-competitive) will always be tempted to adopt alternative practices that have not yet be formally banned, in which case they will expose themselves to some legal uncertainty -- as illustrated by the Michelin saga. In addition, if the presumptions arising from a form-based approach lead to rulings that run counter to the economics of the cases, they will kindle political resistance to competition policy. Political intervention may then take at least some competition policy decisions out of the rule of law. The results then would certainly be less predictable than those of an effects-based approach under the given legal norms.

[4] On this point, see Bolton, P., J. Brodley and M. Riordan, 2000, Predatory Pricing: Strategic Theory and Legal Policy, *Georgetown Law Review*, 88: 2239-2330.

Moreover, without a common foundation in underlying principles, the presumptions as to what forms of behaviour are acceptable and what forms are not may differ from jurisdiction to jurisdiction, depending on each jurisdiction's case law tradition. Such differences impair the predictability of competition policy across jurisdictions. In particular, they are likely to cause frictions in the newly decentralized application of Art. 82 under Regulation 01/2003. A common foundation in underlying principles based on economic theory can help to coordinate and focus the requisite adjustment process and improve the predictability of decentralized antitrust policy under the new Regulation.

Chapter II: Competitive harms

In an effects-based, non-dirigiste approach the analysis of competitive harms naturally focuses on keeping the competitive process open and avoiding the exclusion of actual or potential rivals from the market. In addition, by focussing on the impact of competition policy towards barriers to entry, such an approach guarantees an easier access to markets for new entrants; it therefore contributes to fostering the birth of new activities and firms, in line with the "Lisbon agenda".

In this perspective it seems useful to distinguish three broad typologies of exclusion that differ in respect to the market position of the firms involved and in respect to the specific features that characterize the exclusionary effects: Exclusion within the same market, where an incumbent forces the exit or prevents the entry of a competitor, exclusion in an adjacent market where the dominant firm excludes producers active in markets different but related to its main market, and exclusion in a vertically related market, where exclusion takes places in different stages of the production process.

This chapter therefore presents a framework to deal with competitive harms, organizing exclusionary effects in the three classes distinguished above. Since a given type of exclusion can be implemented through different practices, our classification helps to maintain a consistent approach across practices, uniformly treating all the conduct that can be used to reach a certain type of exclusion. In a given market situation under consideration, the same behaviour can be associated with pro-competitive explanations as well as with exclusionary rationales. In Chapter III we will therefore go through the traditional classification of practices and relate them to the three classes of exclusionary effects highlighted in this chapter as well as to their potential efficiency justifications.

Our treatment is meant to be illustrative rather than exhaustive. Rather than considering all possible effects, abuses, and practices, let alone providing an in-depth analysis of each, we consider some of the most prominent instances in order to show how an effects-based approach makes a difference for competition analysis. Our objective is to elucidate basic principles. We do not try to translate these principles into operational procedures. To do so would require a more comprehensive, systematic and in-depth analysis of the different forms, effects and tradeoffs than we can present here. Presumably such an analysis – and the ensuing operationalization of the basic principles – will be provided when the Commission develops its guidelines for the treatment of abuses under Art. 82.

Section 1: Exclusion within one market

The first kind of competitive harm we consider is the exclusion of a competitor from the market in which the incumbent firm is primarily active. Exclusion can be intended in two different ways: the rival firm can be forced to exit, or alternatively it can refrain from entering the market. Moreover, this category of competitive harm also can include the case when a rival firm, in reaction to the strategies of the incumbent, adopts a passive behaviour and avoids competing fiercely.

The Industrial Organization literature has analyzed a wide set of practices that are consistent with these exclusionary effects. These practices share many common features, which allow us to describe exclusion within one market in general terms.

First, the anticompetitive strategies have a time dimension, and entail an aggressive phase followed by a recoupment period. During the initial phase, the incumbent adopts strategies aimed at reducing the (actual or expected) profitability of the (actual or potential) competitor. For instance, it may reduce the price (predatory pricing) or offer targeted rebates in order to reduce the rival's demand and induce losses;[5] by tying its products the incumbent can prevent entry into the main market by competitors offering complementary products;[6] the dominant firm can use exclusive dealing contracts to deter entry;[7] it may choose a very high level of quality, or offer additional varieties (product proliferation), reducing the competitor's demand or expected profits;[8] it may over-invest in capacity, committing to a low price-high volume strategy;[9] or it may expand its advertising efforts, forcing the rival to increase its promotional expenditures as well.

Although these conducts are focussed on reducing rivals' profits, during this initial phase, in most cases, we also observe an adverse effect on the incumbent's profits, and an improvement in consumer welfare. The reduction of the incumbent's profits comes about due to a distortion in the strategies with respect to the case of normal oligopolistic competition, determining a suboptimal performance in the short-term. For the same reason, the short-term effect on consumers is usually positive, since the rivals' demand and profits can only be reduced if one offers more attractive terms to the customers.

The longer-term effects of exclusionary behaviour move in the opposite direction: once the entrant has been forced to exit (or disciplined into assuming a passive role), or the potential entrant has been discouraged, the incumbent firm can exploit its increased market power recouping the initial losses, while worsening the conditions for consumers.

The overall profitability of the exclusionary strategy hinges on the ability to get rid of the competitors and to prevent further entries into the market. This is difficult to assess both for the incumbent and for an antitrust authority. Antitrust cases involving exclusion can be

[5] See Kreps D., and R. Wilson (1982), Reputation and Imperfect Information, *Journal of Economic Theory,* 27: 253-79, and Milgrom P., and J. Roberts (1982), Predation, Reputation and Entry Deterrence, *Journal of Economic Theory*, 27: 280-312.

[6] See Carlton D. W., and M. Waldman (2002), The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries, *Rand Journal of Economics*, 33: 194-220.

[7] See Aghion P., and P. Bolton (1987), Contracts as Barriers to Entry, *American Economic Review*, 77: 388-401.

[8] See Schmalensee R. (1982), Entry Deterrence in the Ready to Eat Breakfast Cereal Industry, *Bell Journal of Economics*, 9: 305-27, Bonanno G. (1987), Location Choice, Product Proliferation and Entry Deterrence, *Review of Economic Studies*, 54: 37-46, Judd K. (1985), Credible Spatial Preemption, *Rand Journal of Economics*, 16: 153-166..

[9] See Dixit A. (1980), The Role of Investment in Entry Deterrence, *Economic Journal*, 90: 95-106.

opened at different stages of the abusive strategy, and the evidence collected might include only the initial phases of the process. In these cases the competitor, although suffering losses and tighter financial conditions, may still be in the market; and the incumbent has not yet begun to recoup his initial losses. When instead the case is opened after the exclusionary episode has occurred, evidence of actual exclusion and of a subsequent increase in prices can be collected. Hence, recoupment, although essential for the overall evaluation of exclusionary conduct, can be proved with different mixtures of evidence and theoretical elements, depending on the case.

Looking at the related literature, it is interesting for our discussion to focus here on three main scenarios of exclusion, which we may label as reputation, signal jamming and financial predation.[10] Note that, while the rationale for exclusion differs in these three scenarios, many different strategic tools (e.g., pricing, advertising, product differentiation) can be used to implement any of them.

The reputation approach hinges on the competitor's lack of information about the incumbent's incentives to adopt an aggressive market strategy. When the incumbent's incentives are uncertain, the competitor's beliefs play a crucial role in predicting the incumbent's reactions and the profitability of entry. By reacting aggressively early on, the incumbent can then tilt the probability assessment of small or potential competitors (reputation), so as to prevent further entry or to induce exit; the incumbent's profits in the long-run are consequently protected. Hence, even when the dominant firm would have incentives to adopt a less aggressive behaviour in the short-term, the recoupment opportunities offered by reputation can lead the incumbent to adopt an exclusionary strategy.

A second setting where exclusion can be realized involves signal jamming. When small or new competitors have imperfect information on market profitability, a safe strategy would require selective entry into specific market segments. By testing the market the firm can reach a better local knowledge of demand in the neighbourhood of the prevailing prices. In such a situation, the incumbent can reduce its price, improve its quality or launch a local advertising campaign beyond the equilibrium level so as to prevent the competitor from learning the

[10] For a discussion of predation scenarios based on this classification, see, for example Motta, M., (2004), *Competition Policy: Theory and Practice*, Cambridge University Press, Cambridge, U.K.

features of demand in the relevant conditions. Signal jamming can in this way allow the incumbent to delay or deter entry.

Finally, financial predation can be achieved through many different practices that create the preconditions for a negative performance of the competitor. The traditional deep pocket theory of predation simply assumes that the incumbent firm has a financial advantage over the entrant, which allows it to afford the losses of a price war until the competitor is driven out of the market. Today, however, modern Industrial Organization theory challenges this simplistic story, as it does not consider the possibility that an efficient competitor obtains external financing during the predation phase, restoring symmetry between the two firms and making predation unprofitable for the incumbent.[11] A theory of financial predation, consequently, has to explain why this solution does not arise, invoking capital market imperfections that affect the relation between the entrant and its investors. Indeed, even if the incumbent does not make an exclusionary attack, the competitor has to rely on investors, who have a limited ability to monitor the firm's effort, the risk taken, the private benefits extracted, and so forth. This is particular important if the competitor is a new firm. Hence, the financial contract has to provide incentives to induce the firm to repay the investors. Threatening to liquidate the firm or to deny loan extensions in case of insufficient performance are examples of such clauses. Unfortunately, financial contracts that are designed to alleviate agency problems also offer predation opportunities to the incumbent. Aggressive conduct that reduces the cash flow and the profits of the competitor will in fact tighten the conditions for its external finance, reducing the ability of the rival to sustain a prolonged price war. On the other side of this dilemma, any attempt to reduce exposure to predation, for instance by ensuring finance to the competitor even if it performs poorly, would exacerbate the agency problems of the investor.

We have to stress here that, in all the situations described above in which exclusion in one market might occur, most of the strategic variables - e.g., active pricing, product design and tying, exclusive dealing, capacity setting or advertising - that can support an exclusionary purpose can also be adopted when exclusion is not on the agenda and "normal" oligopolistic rivalry prevails. Hence pro-competitive and efficiency arguments can be potentially relevant

[11] See for example Bolton, P,, and D. Scharfstein (1990), A Theory of Predation Based on Agency Problems in Financial Contracting, *American Economic Review*, 80: 93-106.

in any situation involving these practices. For instance, a reduction in prices is a normal reaction to the entry of a competitor, and it can be part of a predatory strategy.

We therefore need to carefully identify the precise story that is supposed to summarize the alleged abusive behaviour, and to compare it with possible alternative explanations, if there are any, that derive from a non-abusive oligopolistic practice. This exercise identifies the relevant elements and the facts that allow discrimination between the lawful and anticompetitive explanations in a given and specific situation.

To illustrate this approach, we come back to the case of financial predation discussed above. Financial predation theory provides a clear setting for antitrust analysis. In order to check if such an explanation may be relevant, we have to consider the following conditions: i) Does the competitor rely on external funding? ii) Do the financing conditions depend on the performance of the borrower? iii) Does the aggressive conduct of the incumbent reduce the ability of the competitor to obtain external finance? iv) Is the impact of the reduced cash flow on the incumbent's financing opportunities limited? v) Is the incumbent is able to recoup the reduced profits once exclusion is realized?

The first two conditions require an analysis of the financial contracts of the competitor; condition iii) does not necessarily imply that the incumbent is pricing below its short-run incremental costs, as it would in other predatory stories, since even an efficient competitor might be in trouble with financial obligations when revenues do not cover all the costs. Finally, condition iv) can be verified by looking at the possibilities for internal financing from other lines of business, or at the impact of localized losses of the incumbent on the volumes of credit received, while point v) can be assessed by considering the prospects of future entry into the market.

To sum up, exclusion within the same market follows a common pattern: a short-run sacrifice in profits and a long-run recoupment of the losses; this pattern appears, for example, in the predation scenarios described above, namely, reputation, signal jamming and financial predation. In each case, exclusion can be achieved through a wide set of strategic tools and practices, which, however, can also be part of normal competition. Hence, in order to distinguish abusive from competitive behaviour we have to carefully identify a precise story of competitive harm and the restrictions on the facts that need to be established in order to substantiate it.

Section 2: Exclusion in adjacent markets

The actions of dominant firms may also generate anticompetitive effects in what are known as related or adjacent markets. An adjacent market is a market horizontally related to the home market,[12] in the sense that competitive conditions in one market depend on competitive conditions in the other, and the products are sold directly to customers. .

The link between the two markets may be established through a variety of practices. It may involve pure bundling, mixed bundling (conditional discounts), tying, full-line forcing, rebate policies, access to interfaces, technical integration, proprietary standards, and compatibility choices. Examples of these practices are the bundling and technical integration of the Windows operating system and the web browser by Microsoft; the tying of nail guns to the sale of exploding cartridges by Hilti; or the early IBM Transamerica case involving the redesign and incompatibility of IBM's CPU with non-IBM peripheral tape drive systems. Such business strategies often yield benefits to consumers, but they may also have anticompetitive effects if the dominant firm distorts competition in the market adjacent to its own home market in order to exclude or discipline rivals in that market, or to influence entry conditions in the home market.

The key condition needed for these actions to cause competitive harm, is that the linkage must place some rivals at a competitive disadvantage so that they cannot compete effectively in the adjacent market or so that they might be deterred from competing in the home market.

In the short-term, some consumers may benefit, while others may be hurt. This is the case, for example, when the linkage is made through bundling, since the design of the bundle will favour consumers with certain preferences and hurt others. Something similar happens if the

[12] This feature distinguishes these markets from vertically related markets (see Rey, P., and J. Tirole (2003), A Primer on Foreclosure, forthcoming in *Handbook of Industrial Organization Vol. 3*, M. Armstrong and R. Porter, eds., Elsevier:Amsterdam).

link between the products is made through technical integration or compatibility choices. In the longer-term, if the linkage hurts the competitive process, all consumers may be hurt due to the exit (or lack of entry) of some rivals. These business practices will be less of a concern whenever several companies engage in providing bundles of complementary goods, since in that instance prices tend to be lower as firms take into account that a low price for a product enhances the demand for its complement.

In addition to having been the subject of potential efficiency explanations, the traditional leverage argument has also been subject to the Chicago critique. According to this view, the dominant firm enjoys a position of market power only in the home market, even if there are two markets and thus two sources of profits to be made. It will have no interest in reducing competition in the adjacent market and will extract all its profit in the home market. As a simple example, consider the case where the two products are complements. Imagine a firm which is dominant in the market for product A and considers leveraging its power in the market for a complement B, where for simplicity we assume that B has no value for the consumer unless consumed together with A. A has value on a stand-alone basis. Under these conditions, the dominant firm will be interested in the provision of B by third parties at the lowest possible cost, since this will boost demand for A. Taking advantage of its dominant position in the home market, the company will be able to increase the price of A and thus benefit from the increased willingness to pay. Under these conditions the dominant company will have no interest in reducing competition in the adjacent market, since any profit increase in market B will be more than offset by lower profits in market A.

The Chicago critique has prompted the development of new economic theories that provide a sound analytical rationale for the existence of strategic leveraging of market power in adjacent markets by dominant firms.

The first line of argument focuses on those cases where the products of the two markets are independent. In this case bundling independent goods can be part of a commitment strategy that tries to prevent entry.[13] By credibly committing to sell the products only as a bundle, the dominant firm commits itself to a very aggressive pricing policy if entry occurs, and this discourages entry. The aggressive pricing is due to the fact that the dominant firm loses sales

[13] See Whinston, M.D. (1990), Tying, Foreclosure and Exclusion, *American Economic Review*, 80: 837-59.

in both the home and the adjacent market when it faces competition in the adjacent market. The commitment story builds on the idea that competing in bundles involves setting lower prices, and this will make entry less attractive.[14] However, this sort of commitment through bundling is not easy to achieve.

Bundling and similar practices that tie or technically integrate related products can also be used as a tool to gain competitive advantage with negative effects on competitors when the products are in principle substitutes, but, nonetheless, there are demand-side benefits from their joint supply. A dominant firm may bundle a product with those in adjacent markets, exploiting these demand-side effects that, as a consequence, will not be available to competitors or entrants.[15]

A third set of arguments deals with markets where the products are complementary, but the Chicago critique does not apply because the dynamics of the market stem from the cost or supply side. For example, consider a market where firms compete through upfront R&D investments and, as a consequence, entry is risky. A potential entrant can enter the market if it succeeds in innovation and obtains a superior technology. By irreversibly tying the two products to one another, the incumbent firm may be able to diminish the expected return in any one market, because successful entry now requires entering both markets simultaneously. Thus, tying makes the prospects of investment less certain, reducing the entrants' incentive for investment and innovation.[16] Or consider a situation where the products are related from the cost side, for example, because the producer may benefit from scale and scope economies.[17] By bundling the two products the incumbent may deny entrants access to a large fraction of the market, and thereby the possibility of achieving minimum efficient scale. A key condition for this result is that entry be easier (or faster) in the adjacent than in the home market. Entry into the adjacent market is just an intermediate step towards competing in the complementary home market. By making entry in the adjacent market impossible, the incumbent tries to prevent entry into his home market.

---

[14] See Matutes, C., and P. Regibeau (1988), Mix and Match: Product Compatibility Without Network Externalities, *Rand Journal of Economics,* 19; 219-234.
[15] See Nalebuff, B. (2003), Bundling, Tying and Portfolio Effects. Part 1. Conceptual issues, DTI Economics Paper Nº1, February.
[16] See Choi, J.P., and C. Stefanadis (2001), Tying, Investment and the Dynamic Leverage Theory, *Rand Journal of Economics*, 32: 52-71.
[17] See Carlton and Waldman (2002), *op. cit*.

In sum, the modern economic analysis of strategic leveraging of a dominant position in adjacent markets has shown that, beyond the Chicago school critique, there may well be sound profit-maximizing reasons why dominant firms attempt to extend their market power beyond their home market. Most importantly, economic analysis provides policy-makers with a fairly detailed roadmap of the specific circumstances that need to be present (such as the relationship between the products, the costs of entry, or the irreversibility of bundling) for these leveraging practices to be judged as detrimental to the competitive process.

Section 3: Exclusion in vertically related markets

The last kind of competitive harm we will consider is the exclusion of a competitor from a market that is vertically related to the market where the incumbent firm is dominant. The common features are as follows: (i) The dominant firm controls a "bottleneck", that is, an input in the production process that is necessary for upstream or downstream firms to exert their activity. (ii) The dominant firm forecloses these vertically related markets by denying or otherwise limiting access to its bottleneck to one or more competitors. For example, the owner of an infrastructure such as rail tracks or port facilities may deny access to this infrastructure to rail transportation service operators, thereby preventing them from providing their services in an effective way. The access denial can be complete, as in the case of a refusal to deal, or partial, as when the bottleneck owner favours some firms (e.g., its subsidiary) to the detriment of others. This foreclosure concern is particularly prominent in industries that have been liberalized through a vertical separation between the infrastructure, featuring the characteristics of a natural monopoly, and the production of goods or the provision of services using that infrastructure. The expected gains from such liberalization, in terms of reduced managerial slack, lower costs and prices, higher rates of innovation, and so forth, could be at risk in the absence of proper access to the infrastructure. This may call for a detailed and on-going regulation of access conditions that falls outside the scope of the present report. In other circumstances, however, a similar concern may arise in unregulated industries.

Foreclosure can be performed in various ways. The bottleneck owner can for example integrate vertically into the targeted market and refuse to deal with potential competitors; airlines' computerized reservation systems have for example involved issues of this type. Relatedly, the "forecloser" may make the bottleneck good incompatible with competitors' products or technologies, or engage in tie-ins and refuse to unbundle. In the presence of economies of scope or scale requiring cooperation among firms in the same market, a dominant group of firms may put its competitors at a disadvantage by refusing to cooperate. In the absence of integration the bottleneck owner can grant exclusivity to a subset of firms, and thus de facto exclude their rivals. Alternatively, it can favour some competitors over others; it can for example discriminate explicitly, e.g., through personalized rebates, or implicitly, e.g., through loyalty programs or growth-based rebates that are formally available to all but tailored to the needs of specific customers. Similarly, substantial quantity discounts may allow the survival of only a few customers; for instance, a large enough fixed fee can transform a potentially competitive downstream industry into a natural monopoly industry. Such considerations (besides many others) played a role in the process of enacting the Robinson-Patman Act in the US in 1936.

The traditional foreclosure concern is that the bottleneck owner may leverage its market power in the related markets. However, as pointed out by the Chicago School critique, there is a single final market and therefore only one profit to be reaped. And since the dominant firm can get this profit by exerting its market power over the bottleneck, it has no incentive to distort competition in the other markets; on the contrary, imperfect competition in the downstream market may actually create distortions and reduce the profitability of the bottleneck, e.g., by reducing the variety or quality of the goods and services produced.

While the Chicago critique is correct, anticompetitive effects may still arise in specific circumstances.[18] For example, as already mentioned in the case of adjacent markets, the bottleneck owner may deter competition in a vertically related market to protect its home market.[19] Alternatively, the bottleneck owner may face a commitment problem, which makes it difficult to exert its monopoly power without engaging in exclusionary practices. Indeed, once it has sold access to a first competitor, it has an incentive to provide access to other

18 See Rey and Tirole (2003), *op. cit.*, for an overview
19 See Carlton and Waldman (2002), *op. cit.*

competitors as well, even though those firms will compete with the first one and reduce its profits; this opportunistic behaviour reduces ex ante the bottleneck owner's profit (in the example just given, the first firm is willing to pay and buy less); more generally, the bottleneck owner would like to commit to a certain volume of access, so as to limit competition and profit dissipation, but it may be tempted to grant more access when dealing bilaterally with each competitor; as a result, competition in related markets "percolates" in the bottleneck market and dissipates the dominant firm' profit. When this commitment problem is serious and cannot be solved through adequate contractual provisions, the bottleneck owner may wish to restrict or eliminate competition in these related markets through the types of practices mentioned above.[20] For example, refusing to deal with all downstream firms but one, or entering into an exclusive dealing agreement with that particular firm, eliminates downstream competition and thus fosters the upstream firm's ability to exploit its market power. Remarkably, banning discrimination would also help the bottleneck owner to resist demands for selective price cuts and thus contribute to maintaining high prices. Vertical integration also constitutes an alternative solution to the upstream firm's commitment problem.

In such situations, the intervention of competition authorities may generate more competition in the related markets and thus in the industry as a whole. While any such intervention benefits consumers, e.g. in the short-run through lower prices or in the long-run through higher rates of innovation in the related markets, it also regulates the bottleneck owner's rate of return. In the long-run it may thus have an adverse impact on the dominant firm's incentives to invest or innovate and may for example impede the development of a key infrastructure. No prospective licensee would want to pay much for the use of a new technology if it knew that the licensor would "flood the market" with similar licensees; mandating access through additional licenses would thus reduce the innovator's profitability and consequently its incentives to invest in R&D.

To sum up, modern economic analysis has gone beyond the Chicago school critique and identified several reasons why a firm may use its dominant position in one market to distort competition in vertically related markets. Furthermore, several alternative types of practices

---

20 See Hart, O., and J. Tirole (1990), Vertical Integration and Market Foreclosure, *Brookings Papers on Economic Activity (Microeconomics)* 205-285.

can be used to that purpose – e.g., refusal to deal, exclusive dealing, prohibitively high access prices for downstream rivals, etc. Therefore, when vertical foreclosure is the concern, one should treat these alternative practices in a consistent manner. The economic approach also emphasizes the relation between ex post intervention and ex ante investment incentives. It can thus help competition authorities to identify the specific circumstances calling for intervention.

Chapter III: Implications for practices

The previous chapter has highlighted the economic approach to different types of competitive harm. We now discuss how this approach can be put to work in Article 82 cases. The above discussion already stressed that several alternative practices can often serve the same anticompetitive purpose. Conversely, when a complaint arises or when a competition authority suspects an abuse of dominance, from an economic perspective the first question is: What is the nature of the competitive harm involved in that case? Pursuing this question brings several benefits.

First, it allows a clear identification of the "economic toolbox" of relevant arguments. For example, the type of reasoning involved for adjacent markets is quite distinct from those that may be relevant when exclusion arises in the dominant firm's core market. Second, while creative imagination in business relations can lead to an infinite number of different practices, there are not that many types of competitive harm and for each one the established toolbox of relevant, consistent economic arguments is relatively limited; in addition, each line of reasoning outlines the key facts that need to be checked. Thus, identifying the nature of the competitive harm at stake can both facilitate and speed up the investigation process, and contribute to maintaining high standards of predictability. Third, as already noted this approach guarantees a consistent treatment of the alternative practices that could serve the same anticompetitive purpose.

Once a competitive harm has been identified and the relevant facts established, the next step should be to see if pro-competitive effects might counterbalance them. There again, an economic approach first identifies the nature of the benefit for competition and the facts that need to be established. It is only after these steps that a proper balance can be assessed.

In this chapter we illustrate this approach for various practices. We start with a general discussion of price discrimination, which is relevant for many other practices: rebates, for example, typically involve some discrimination, while exclusive dealing can be interpreted as an extreme case of discrimination; likewise, tying and predatory pricing often involve some form of discrimination. While price discrimination can be part of an exclusionary strategy, it often can also bring benefits, particularly when large fixed costs are present; in a given price discrimination case, it is thus again important to identify the exact nature of the potential anticompetitive harm, as well as the possible pro-competitive effect of price discrimination in the context of the case.

We then discuss several classical types of practices: rebates, tying, refusals to deal, exclusive dealing and predatory pricing. In each case, we first relate the practice in question to the exclusionary effects outlined in Chapter II. We then review (some of) the potential pro-competitive benefits of the practice; finally, we illustrate, using an example, how the effect-based approach could be applied to the practice.

Section 1: Price Discrimination

Price discrimination consists in charging different prices for different units and/or to different customers.[21] It can take various forms. In particular, it can be explicit, as when different customers are offered different prices on the basis of their age (e.g., reduced prices for children or senior citizens) or on the basis of other characteristics (e.g., student fares,

[21] This simple definition raises tricky issues when, for example, different customers involve different costs. Should "no discrimination" then mean the same price for all, or should cost differentials be borne by the customers? Interestingly, while competition law usually insists that any difference in prices should reflect a cost difference, universal service obligations, for example, tend instead to insist on "equal treatment", according to which all customers should be offered the same price in spite of possibly large cost differentials.

geographical segmentation), or it can be implicit, as when all customers are formally offered the same menu of options, but different customers de facto pick different options and thus end up paying different prices (thus, for example, volume discounts allow firms to offer better deals to large customers).

Antitrust policy and jurisdiction have traditionally been very strict about price discrimination, sometimes even treating it as a *per se* offense against the law. Price was – and is – seen as a tool by which the dominant firm exploits its power to earn more profits. Price discrimination is also considered to be unfair because some people have to pay more for the good in question than others. This latter point, in particular, resonates in public discussions about pricing.

Economic analysis has also long put a lot of weight on the exploitative effects of price discrimination allowing the dominant firm to earn more profits. Economic analysis has also stressed that the distribution of output across consumers tends to be inefficient if different consumers pay different prices and presumably put different valuations on the last units they purchase. These arguments imply that any price discrimination which reduces (or barely increases) total output is necessarily detrimental for total welfare and even more so for consumer welfare.

However, more recently economic analysis has also shown that, in some circumstances, price discrimination can increase total welfare and even consumer welfare. In particular, this is likely to be the case if price discrimination permits a significant expansion of output. [22] This might happen, for instance, because additional offerings at lower prices permit the firm to serve additional customer segments. Moreover, if the firm has significant sunk investments, the greater profits which the firm obtains from price discrimination may be necessary to provide a return on these investments. These returns may also encourage the firm to invest more, providing additional pro-competitive effects in the future. In all these cases, price discrimination is likely to benefit consumers, sometimes even those who pay the higher prices: Even though they pay more than other consumers, they benefit from the fact that the firm invests and makes its output available to them.

[22] For an extensive welfare analysis along these lines, see for example Varian, H., (1989), Price Discrimination, in R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*, North-Holland, Amsterdam, 597-654.

Such considerations have led economists to be skeptical about using the simple notion of “fairness” or “unfairness” to assess price discrimination. In the examples given, prohibiting price discrimination on the grounds of unfairness to those consumers who have to pay a higher price may end up making these very consumers worse off.

At the same time, one must worry about the fact that any assessment of the fairness or unfairness of price discrimination is necessarily based on a *local* assessment of distributive effects. There is no way of telling whether such a local assessment is consistent with global concerns for distribution.[23] With price discrimination, consumers that are more sensitive to prices will be offered lower prices and will thus benefit while the others are more likely to be hurt. From a global perspective, the assessment of this distributive effect depends, for example, on whether the greater sensitivity to prices reflects a higher level of education or a greater need to turn each euro around twice before spending it. In the first case, the distributive effect favours people with a higher education level, in the second case, people who are poorer. Is a pricing scheme that favours poorer people really “unfair”? However, such global distributive concerns can hardly be made the subject of antitrust proceedings under a rule of law. Therefore, it seems advisable to assess price discrimination less in terms of fairness and more in terms of pro-competitive and anti-competitive effects.

For price discrimination does also have an impact on the way firms compete. Thus, price discrimination tends to intensify competition among oligopolists and thereby raise consumer surplus at the expense of industry profits. Consider, for example, a case of customer poaching, where customers face switching costs and firms offer lower prices to the customers belonging to their rivals’ customer segments.[24] Basically, such discrimination intensifies competition, because it makes it possible for a firm to attack its rivals’ customer bases, as well as new customer segments, while maintaining higher margins on its own installed base. But since all the firms have similar strategic incentives to exploit price discrimination, the industry faces a prisoner’s dilemma situation, and competition is more intense than with uniform prices. More

[23] For the distinction between local justice and global justice, see Elster, J. (1992), *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*, New York & Newbury Park: Sage.

[24] For detailed analytical models, see, for example, Fudenberg, D., and J. Tirole, (2000), Customer Poaching and Brand Switching, *Rand Journal of Economics*, 31: 634-657, or Chen, Y. (1997), Paying Customers to Switch, *Journal of Economics and Management Strategy*, 6: 877-897; for a borad perspective, see Stole, L. (2005), Price Discrimination and Imperfect Competition, forthcoming in *Handbook of Industrial Organization Vol. 3*, , M. Armstrong and R. Porter, eds., Elsevier:Amsterdam.

generally, price discrimination leads to competition on a customer per customer basis, which is likely to be more intense than when firms are constrained to offer the same price to all customers. Price discrimination encourages the firm to target more customers, by allowing the firm to offer specific deals to these customers without compromising the profits achieved on more captive customers. In particular, this reasoning militates for allowing an incumbent to respond non-uniformly to limited entry ("meeting competition" strategies).

Even when a firm has enough market power to be dominant, the ability to engage in price discrimination might very well reduce its market power. For example, the firm may be tempted to grant concessions when a particularly good customer insists on getting a better deal, or when a customer suffering from a temporary downturn asks for a specific treatment. And indeed, the ability to negotiate good deals in bilateral bargaining constitutes one the drivers of competition, particularly when the customers are themselves competing against each other in a downstream market. In such circumstances, a ban on discrimination actually helps the dominant firm resist requests for lower prices and thus exploit its market power. Indeed, if any dominant firm is obliged to offer similar deals to all equivalent consumers, it can no longer offer customer-specific discounts based on individual bargaining. But then it will be more reluctant to grant a concession to a particular customer if the concession must be generalized to all other customers. Thus, a no-discrimination requirement serves as a very effective device to enhance the market power of the dominant firm, relative to a situation where price discrimination, in the form of customer-specific bargaining, would be allowed.[25]

In many important applications it is not sufficient to evaluate the welfare effects of price discrimination by comparing prices, production, and consumer surplus to those associated with uniform pricing within the framework of a given market structure. The option of price discrimination may affect entry (or exit) decisions and long-run investment decisions in ways that may significantly affect the long-run performance of the industry. For instance, the use of certain types of rebates warrants detailed analysis, as will be discussed in the next section.

Overall, these arguments thus support a view according to which price discrimination may, in fact, promote efficiency and benefit consumers even when firms have significant market

[25] The impact of a non-discrimination requirement here is similar to the effect of "most-favoured customer" clauses that allow a firm to commit itself to maintain high prices.

power. Price discrimination can, however, also serve anticompetitive purposes. In particular, it can be associated with any of the exclusionary effects highlighted in the previous chapter. Selective price cuts can for example reduce the cost of predatory pricing strategies to deter entry into the dominant firm's home market or into adjacent markets (in the latter case, mixed bundling also involves some form of discrimination); and access price discrimination can be at the heart of vertical foreclosure cases. But then, it is the exclusionary effect that causes the anticompetitive harm, rather than discrimination per se. This therefore again calls for an effects-based approach, which first identifies the type of exclusionary effect involved, rather than a form-based approach, to the spectrum of business practices in which price discrimination is a central element.

Section 2: Rebates

Rebates may come in the form of pure volume discounts: the firm then offers a rebate if the quantity bought by a customer exceeds a given threshold. In this case, the rebate may apply to the incremental quantities (only those above the threshold) or to all the units bought by the customer as soon as the threshold is reached. In the former case, the rebate scheme induces a progressive discount, and in all cases, we have (possibly discontinuous) non-linear pricing. By contrast, when it applies to the whole quantities bought by the customer, it amounts to switching to a new price scheme as soon as the threshold is achieved. Rebates can also be offered to a customer whose *growth* in the volume of purchases lies above a given threshold. Rebates on a particular product can also take the form of a more advantageous offer to customers who buy another product together with the initial one. Rebates then involve mixed bundling (see chapter II, 2 and chapter III, 3). Finally, fidelity rebates may be offered: for instance, rebates can be conditional on the client buying all its quantities, or at least a given percentage of them, from the firm. In most cases, rebates imply some form of discrimination between customers. To this extent a part of the analysis of Section 1 (in this chapter) above is relevant.

a) Potential anticompetitive effects

Generally speaking, a consequence of this practice is that competition operates on a customer basis rather than on a "unit sale" basis, which can be anti- or pro-competitive, depending on the circumstances.

The anticompetitive effects of rebates may be of any kind described in Chapter 2:

- Rebates can exclude actual or potential competitors from the market on which the firm is dominant. This is the case, for instance, for selective rebates offered to those of the customers of the firm that would switch to a new entrant were the rebate not offered, or if the rebate is conditional on the percentage of quantities bought by the customer from the firm. In most of these cases, rebates may be associated with predatory pricing on some of the units sold. Moreover, like predatory prices, rebates induce short-run sacrifices and may have exclusionary effects either by inducing exit or by discouraging entry.
- They can also involve horizontal foreclosure: this is for instance the case if the rebate is offered to a customer that buys the product of the dominant firm on an adjacent market together with the product on the main market. This is an example where rebates also tie together the products on two different markets.
- Finally, rebates can induce vertical foreclosure. This is the case for instance when a producer offers to its retailers in order to discourage them from selling competitors' products. They may then be associated with an exclusivity clause. Competitors are therefore unable to obtain access to a distribution network to sell their products. Of course, this mechanism may appear in any vertically related market, where one of the stages plays the role of an essential facility. In these "essential facility cases", rebates may eliminate downstream or upstream competition in order to better exploit upstream or downstream market power.

b) Pro-competitive effects and efficiency considerations

Efficiencies may be a cause or a consequence of rebates. A general way of assessing the dominance of pro-competitive effects over anticompetitive ones is to check whether total output has increased or not. Pro-competitive effects of rebates may for example appear in the following circumstances.

Since rebates allow high and low demand elasticity consumers to be treated differently, elastic demand segments tend then to generate lower margins. Consumers with a high elasticity of demand thus benefit from the practice, although consumers with a low elasticity may suffer from it; the overall effect on consumer welfare is thus a priori ambiguous. But in the spirit of Ramsey pricing, in this way rebates may also allow for the recovery of fixed costs, and thus encourage R&D investments that involve such large fixed costs. As a result, rebates are more likely to have a pro-competitive effect when high fixed costs are involved.

Rebates that are targeted to those consumers who are more likely to switch to competitors imply a more intense competition for these consumers; they clearly benefit from this situation. The other consumers may indirectly benefit from an increased pressure on the price they face. Moreover, prohibiting selective rebates as a reaction against competitive pressure may constitute excessive interventionism in the competitive strategies of firms on the part of competition authorities.

In a vertical relationship in particular, rebates that take the form of non-linear pricing may be used as an incentive mechanism to induce efficient behaviour of retailers. For example, rebates can be used to increase retail margins on additional volumes, so as to encourage retailers to promote the product. While a uniform reduction in the wholesale price might have the same impact on retailers' incentives, it would be more costly for the supplier. Hence, rebates allow suppliers to provide incentives at a lower cost, thereby encouraging suppliers to provide more incentives and thus to compete more intensively. More generally, rebate schemes can enhance efficiency by solving adverse selection or moral hazard problems.

Rebates may also generate efficiency gains for the dominant firm, for instance, economies of scale for this firm, or economies of transaction costs for the customers (the buyer concentrates its purchase on a single seller).

These elements lead to a general intuition: rebates that take the form of pure quantity rebates are more likely to be motivated by efficiency considerations than fidelity rebates. It is, however, difficult to demonstrate that efficiency considerations motivate the rebate scheme. In particular, the mere form of the rebate may not constitute a clear indicator; for instance, efficiency considerations might require personalized rebate schemes, tailored to the “size” of the retailer, which could take the simple form of “market-share” fidelity rebates.

c) Implementation: An example

To deal with an alleged story of anticompetitive rebates, the competition authority should follow the analysis that derives from the previous considerations, e.g., it should first identify the kind of exclusionary strategy at work, and look for possible pro-competitive effects of the practice. This approach clearly stands in contrast with a *per se* prohibition; instead, it involves a sound analytical framework and relevant data. We illustrate this approach with reference to a particular example.

The view of rebates that prevails in the contemporaneous approaches recognizes that the anticompetitive effects of this practice can indeed dominate, and that in some cases rebates should be prohibited. For example, a supplier may use rebates so as to impose a penalty on new entrants. A customer will switch to a new supplier only if the latter offers a price that is lower than the price charged by the incumbent supplier minus the rebate.[26] Thus, the rebate is analogous to a penalty paid by the entrant; it plays the role of an entry fee, designed to extract some of the efficiency gains of new entrants, and by the same token it creates a barrier to entry. The rebate clause thus imposes an external effect on potential entrants, and it is this externality that makes it profitable for the incumbent supplier and the customer to enter in this type of arrangement. Note that the entrant will enter only if its costs are so low that entry remains profitable despite the entry fee generated by the rebate. Consumers are harmed because the probability of entry is reduced and prices are raised.[27]

In such a case the competition authority should proceed by asking the following questions:

[26] Indeed, granting a fidelity rebate of *10* on a reference price of *100* could also be presented as a “penalty for breach” of the same amount, based on a reference price of *90*; in both cases, to attract the customer a new supplier will have to compensate the customer for the lost rebate, and the provision thus acts as a “tax” on the new supplier

[27] Cf Aghion and Bolton (1987), *op. cit*. This model, designed to deal with exclusive dealing, suits perfectly well the analysis of rebates. This shows again that various practices may achieve the same anticompetitive effect.

- What exactly is the negative externality imposed on a third party that explains why the buyer agrees to accept an anticompetitive exclusive dealing contract? If the competition authority cannot identify such an external effect, then there is no anticompetitive effect and the case should be dismissed.
- What is the magnitude of the penalty imposed on the entrant? This can be approximated by computing the reduction in price that the entrant should offer in order to be able to enter the market at various quantity levels.
- Were in fact efficient competitors kept out of the market by the penalty imposed on them?
- Efficiency effects have to be carefully assessed; among the critical factors are the ability of downstream firms to pass on a reduction in their own input prices and the incentives given to these downstream firms to exert effort.

Section 3: Tying and bundling

a) *Tying and bundling may take various forms*

First of all, tying consists in making the purchase of one good (the tying good) conditional upon the purchase of another one (the tied good), whereas bundling refers to the sale of two products together. Bundling may be pure (the goods are available only together) or mixed (they are also available separately). Therefore, when tying is at work, the tied product may be bought alone, whereas when pure bundling is at work this is impossible. Tying and bundling can be achieved through technological links or through a contractual practice. Tying and bundling may concern goods that are complements, or substitutes, or independent. In the latter case the only link between markets is the fact that both goods are bought by the same consumers. For the sake of exposition we will focus on pure bundling, but the analysis also applies to mixed bundling, although potential anticompetitive effects are likely to be less stringent in that case. In addition, mixed bundling has another advantage, since it makes it possible to offer multiple price formulas adapted to the needs of different consumers. Consumers who want only one of the goods can still buy it, while those who want to buy both face lower prices for the bundle. Therefore, as in the above discussion of price discrimination, mixed bundling may induce higher consumption from some consumers.

The problem is to identify cases where tying is anticompetitive, that is, profitable for the firm that implements the practice, while inducing exclusion and hurting consumers.[28] These cases are relatively scarce, a condition which does not favour a prohibition *per se*.

b) Tying and bundling can give rise to many anticompetitive effects.

- Bundling may serve an exclusionary purpose on the tied market.[29] This is the commonly expected effect of bundling: since the consumer buys both good A and also good B from the dominant firm, either necessarily (if there is pure bundling) or because it is advantageous for him (in the case of mixed bundling), a competitor on market B cannot profitably sell its product, even if this a better quality product or produced at a lower cost. This effect is analogous to predatory pricing. Bundling may also be associated with fidelity rebates tying products A and B together.
- Bundling may be used in order to protect the dominant firm's home market. This is the case, for instance, when it is easier for a competitor to enter market A if it is active on the market B with a complementary good. Then, in order to deter entry of its competitor on market B into its home market A, a firm might want to sell bundles of its own products, therefore eliminating its competitor on market B, and making entry into market A unprofitable.

In the last case, bundling forces competition between bundles rather than between elementary goods. This form of competition is not necessarily a bad thing, as it may well be more intense (see below), but it may force the rival to enter several markets simultaneously, which may make entry costlier and/or riskier.

[28] Chapter 2 recalls the Chicago arguments concerning tying and bundling.

29 This is also the common "leverage effect" that is feared in mergers involving firms active on adjacent markets.

c) Pro-competitive effects of tying and bundling

Of course, the analysis of the overall effect on consumers has to take into account the fact that the strategies of the dominant firm linking the two markets may also be the result, not of an attempt to exclude competitors, but rather of an attempt to improve the efficiency and quality of the products supplied in the market.

Like rebates, tying may allow for the recovery of high fixed cost - even when goods are independent. In this case, consumers might also benefit from bundling. When goods are complementary, additional arguments are at work. Bundling also reduces transaction costs for the consumers, who would, in the case of mandatory unbundling, be forced to buy all the components separately; still in the case of complementary products, tying guarantees that two components of a system good are compatible. Bundling may also boost the demand for the complementary good when it is associated with a low price (even zero) for the tied good. In the case of complementary goods used in variable proportions, tying may also serve as a metering device for the tying good: a well known example is that of photocopiers sold with the paper or with after-sales services. The consumption of the latter reveals the intensity of preferences for the former and this information allows the firm to discriminate between consumers, which may be beneficial to some of them. Making the customers pay for usage as well also allows the firm to lower the price of the equipment. This can reduce distortions in replacement decisions (both by inducing consumers to internalize the cost of usage and by distributing mark-ups over both acquisition and usage, rather than placing them on acquisition only); this can also be used to credibly convince customers to try and switch to new types of equipment.

Efficiency justifications are particularly common when the linkages are established by technological means (technical integration, standards, and compatibility): costs savings, quality improvements and the overcoming of information problems are important sources of gains. For example, when the quality of two complementary products is uncertain, supplying them jointly may alleviate this informational problem. Similarly, bundling complementary products may reduce the inefficiency created by double marginalization: the standard "Cournot argument", according to which vertical integration prevents the addition of multiple mark-ups, also applies to any two complementary products, not just to vertically related ones.

Indeed, tying leads to "system-based" rather than to "component-based" competition, in which it is easier to internalize the effect of a price cut for one component on the sales of the system. As a result, this system-based competition is likely to be more intense.[30]

To sum up, depending on the nature of competition, the cost structure on the tied market, the magnitude of costs savings associated with bundling, and the existence of strategic reasons, bundling can have exclusionary effects or pro-competitive effects and should thus be analyzed in the light of the effects of the practice. The potential for efficiency gains is more limited when the linkage is achieved through pricing schemes and bundling than when it is achieved through technological integration. Still, as in the case where it is used as a metering device, bundling can enhance welfare.

d) An example

The naive view of (pure) bundling is the following: a firm that is dominant on market A fears entry from a competitor present on market B, where the dominant firm offers its own good. Goods A and B are complementary. Therefore the dominant firm offers (pure) bundles of A and B and prevents consumers from buying the competitors product. This protects the dominant firm from the competitor's entry into its home market. . However, this is the case where the Chicago critique fully applies, as shown in chapter II.2: the firm has no interest in reducing competition on market B. Since there is only one monopoly profit to be made, it would rather allow the competitor to be active on market B and take advantage of an increased demand on market A in order to enjoy a monopoly margin on each unit of good A. In the more sophisticated version of this argument,[31] entry into one market depends on the success of entry into another market: a competitor can, for example, enter market A only if it can first enter market B with a better or cheaper product than that of the dominant firm. The entry strategy thus involves a dynamic scenario, in which first the dominant firm faces a more efficient competitor on market B, and later the same or another firm can enter market A. The dominant firm has an incentive to sell a bundle of A and B: in this way it can deter the first competitor in market B, and entry into market A is then less likely.

[30] In the absence of tying, cutting the price of one component also benefits mixed systems, where some of the components are sold by other firms; as a result, the incentive to cut prices is lower and competition is less intense; see Matutes and Regibeau (1988), *op. cit.*
[31] See Carlton and Waldman (2002), *op. cit.*

There again, there is a commitment problem: once the competitor has entered the market, the dominant firm would have an interest to sell goods A and B separately. Technological bundling may however be used as a commitment device. As was recently highlighted by J. Tirole, the bundling strategy involves sacrifices with regard to a strategy where the firm would accommodate entry, as in predatory pricing strategies.[32]

Again, following an effects-based approach regarding this practice requires that the competition authority first build a consistent story showing what kind of exclusionary effects are at work and that it check the facts that would allow for a pro-competitive explanation of the practice. In our particular case, the competition authority should here check the following points:

- Does the situation present dynamic features (sequential entry)?
- Does the existence of a competitor for the complementary product make the entry into market A more likely (either from the competitor present on market B, or by a third firm that finds it profitable to have a competitive market for good B)?
- Is the quality of the good offered by the rival on market B higher than that of the dominant firm?
- Is the bundling practice credible? For instance, is it achieved through commercial or through technological bundling?

In addition, the competition authority should of course check to ensure that the possible efficiencies of the practice are not the major determinant of the firm's behaviour. This is particularly important where compatibility problems are involved (see above).

[32] See Tirole, J. (2005), A Primer on Tying, *Competition Policy International*, 1: 1-25.

Section 4: Refusal to Deal

Refusal to deal may take several forms. The dominant firm may simply refuse to supply the essential good, or it may charge a prohibitively high price. It may make the bottleneck good incompatible with the products offered by its competitors, or it may tie the essential good to some other good, thus making it unattractive for competing firms to buy the bundle. Or it may sign an exclusive dealing contract with one particular firm, thus excluding others from the market (see Section III.5 below).

a) Potential anticompetitive effects

Refusal to deal is a typical form of exclusion in vertically related markets when a dominant firm controls an input that is essential for production by competing firms in a downstream (or upstream) market. By refusing access to the input good the dominant firm extends its monopoly from the market for the essential input to the potentially competitive downstream market. However, the monopolization of a downstream market need not have anticompetitive effects per se. After all, there is only one final market and only one monopoly profit to be reaped. If the dominant firm is able to capture the monopoly profit of the final market even if there is downstream competition, then monopolization of the downstream market cannot have anticompetitive effects because there is no competition anyway. In this case a refusal to deal with downstream firms is likely to be motivated by efficiency arguments.

The dominant firm may also be a group of firms or an industry association refusing access to a jointly owned facility. In this case a competitor is excluded on the same or on a horizontally adjacent market. Famous examples include *Associated Press (1945)* and *Aspen (1985). In the latter case,* where a three-mountain ski resort refused to make lift tickets available to a competing one-mountain ski resort.[33]

[33] See e.g. Ahern, P. (1994), Refusal to Deal after Aspen, *Antitrust Law Journal*, 63: 153 ff:

b) Pro-competitive effects and efficiency considerations

Upstream firms may be worried that their reputation will suffer if their product is offered by inferior downstream firms. If the upstream firm cannot effectively monitor and control the downstream firms, excluding them from the market may be the only possible way to increase consumer surplus and industry profits. Similarly, for technological reasons, it may be necessary to closely monitor downstream production by the upstream supplier. This may only be feasible in a vertically integrated firm. Or downstream firms may free-ride on the marketing expenses of the upstream firm. In order to recoup this investment, the dominant firm may have to exclude downstream competition.[34]

But even if a refusal to deal harms consumers in the short-run, it may be socially beneficial in the long-run. If the bottleneck is the result of investment or innovation activities of the dominant firm then forcing the firm to give its competitors access to the bottleneck is an expropriation of the returns of the firm's efforts. This may discourage this and other firms from investing in the future, and it may reduce the incentives to innovate. Tolerating a (temporary) monopoly may be the best way to promote investment and innovation incentives and thus dynamic efficiency.

If the bottleneck is due to an intellectual property right, the competition authorities should be particularly reluctant to interfere. Intellectual property rights have been granted by the state in order to create market power and to give innovators a reward for their efforts. Thus, it is inconsistent if the state interferes with these rights ex post and takes market power away. Indeed the mere prospect of interference affects the parties' bargaining powers in negotiating a voluntary agreement. To the extent that the rival firm obtains favourable terms by threatening to sue in order to impose a duty to deal, the bottleneck owner obtains a lower return on his investment. Because there is no active intervention, the effect is not visible, but nonetheless it reduces the incentive to innovate.

[34] For a more detailed discussion of efficiency defenses for vertical foreclosure, see Rey and Tirole (2003), *op. cit.*, Section 5.

c) *Implementation: an example*

Refusal to deal increases the market power of a dominant firm only if it was unable to fully exploit its monopoly power over the bottleneck good beforehand. For example, this is the case if the dominant firm has a problem committing to charging all the downstream firms the monopoly price. The reason is that once the monopolist has contracted with one downstream firm on the supply of the essential good, he has an incentive to supply the other firms at more favourable conditions in order to further increase his profits at the expense of the first downstream firm, which then has to compete with the other firms on the downstream market. However, the downstream firms anticipate such opportunistic behaviour by the upstream firm and will buy the essential good only at a discount. This reduces the profits of the monopolist. He may restore his monopoly power only if he manages to eliminate competition on the downstream market altogether. In such a case, refusal to deal could have anticompetitive effects.

If the competition authority suspects that such a mechanism is a work, it should proceed as follows:

- First it has to establish that there are anticompetitive effects. Suppose that the dominant firm sold the essential good to downstream firms in the past, but that it now refuses to deal with them. If the price for the final good remains unchanged and if the stock price of the downstream firms is not affected, then it is unlikely that the refusal to supply has increased the dominant firm's market power and reduced social welfare. If the dominant firm can come up with a convincing efficiency defence for the refusal to deal, the case should be dismissed. On the other hand, if the dominant firm found it difficult to commit to the monopoly price when there was downstream competition and if it manages to raise the final price paid by consumers by monopolizing the downstream market, then this is an indication of anticompetitive effects.
- What is the source of the bottleneck? If the bottleneck is mainly due to the investment and innovation efforts of the dominant firm, the returns of this investment should not be expropriated and the competition authority should not interfere even at the cost of a static inefficiency (a temporary monopoly). The competition authority should be particularly reluctant to interfere when the source of the bottleneck is an intellectual property right.

However, if the bottleneck stems from historical legacy, economies of scale or scope, or network externalities, an intervention may be justified.

- Is the intervention likely to be effective without impairing efficiency? Enforcement may be difficult and costly. Moreover, the competition authority is likely to be drawn into the process of determining the *terms* on which the dealing must take place, i.e. prices, conditions and technical specifications. The authority is not really qualified to set such terms, so its intervention may cause substantial inefficiencies. Thus, the competition authority should be aware of the harm that it may cause, and intervene cautiously, refraining from active involvement in the dealing terms.[35] Structural remedies, such as divestitures and line-of-business restrictions, often involve substantial transaction costs and should be considered only as a last resort.

The competition authority should also be aware that its approach to refusal-to-deal cases affects outcomes even when firms reach a voluntary agreement so, in fact, there is no refusal to deal. The possibility that the bottleneck owner may be sued affects the participants' bargaining positions. If the rival firm can threaten to sue in order to impose a deal, it is in a much stronger position than if the owner of the bottleneck can refuse to deal. As long as these cases are not contested in legal proceedings, these effects are not visible. Nevertheless, they raise the same concerns as the authority's handling of refusal-to-deal cases itself: If the anticipation of strict policy intervention leads to a voluntary agreement at low access prices, this reduces the returns to the bottleneck owner's investments; if the anticipation of weak policy intervention leads to a voluntary agreement at high access prices, this restrains the rival's ability to compete downstream. To avoid competitive harm from these effects of anticipations, the competition authority should have clear guidelines for the assessment of refusal-to-deal cases, providing well-specified standards by which to compare exclusionary concerns and concerns about returns on investments.

[35] In this context, it is noteworthy that in the United States court interventions under the "essential facilities" doctrine have usually left the determination of terms of dealing to the parties.

Section 5: Exclusive Dealing

Exclusive dealing refers to all practices that commit a firm to deal exclusively with some vertically related firms but not with others. For example, a dominant seller may prohibit buyers from dealing with its competitors, or a dominant seller may commit to deal exclusively with one (or several, as in the Eurotunnel case) vertically related firm (or firms), but not with others.

a) Potential ant-competitive effects

Exclusive dealing may be used against a (potential) rival on the same or on a horizontally adjacent market if enough buyers agree to exclusively buy the good from one dominant firm and if the rival finds it more difficult to enter this or the adjacent market should he be prevented from dealing with these buyers. Exclusive dealing contracts may also restrict entry to a vertically related market if the dominant firm is the sole producer of a bottleneck good that is essential for a downstream market and if the dominant firm commits to supply the bottleneck good exclusively to one firm but not to others. In this case exclusive dealing may be used as a substitute for vertical integration and have similar effects to a refusal to deal (see Section III.4).

It is often argued that exclusive dealing arrangements cause competitive harm because they raise the cost of entry into the market or deter entry altogether. However, this argument is too simplistic. A rational buyer would not be willing to sign an exclusive dealing arrangement if such a contract obliged him to buy from an incumbent while a (possibly more efficient) competitor is willing to enter the industry. An exclusive dealing contract will be agreed to only if it is beneficial for both parties. Thus, the exclusive dealing contract must either be efficiency enhancing, in which case there is no competitive harm. Or it must increase the payoffs of the two contracting parties at the expense of a third party. In recent years several economic theories have been developed showing that exclusive dealing may indeed have anticompetitive effects if it imposes an externality on a third party. For example, an exclusive dealing contract that makes entry more difficult may be used to extract rents from a potential entrant.[36] Alternatively, if the buyers cannot coordinate their actions then the seller can bribe some buyers into an exclusive dealing arrangement at the expense of the other buyers.[37] The common denominator in all of these anticompetitive effects is that the exclusive dealing contract imposes an externality on third parties.

b) Pro-competitive effects and efficiency considerations

There are several efficiency defences for exclusive dealing arrangements. The most frequent argument is that exclusivity may be necessary to protect and encourage relationship-specific investments. For example, a manufacturer will invest in the training and education of downstream firms only if it can be assured that this investment will not be used to benefit upstream competitors. Thus, in order to protect its investment, the manufacturer may have to insist on an exclusive arrangement. Similarly, a manufacturer may adopt an exclusivity clause in order to promote "retailer loyalty", i.e. to encourage the retailer to tailor its promotional efforts towards the manufacturer's product.

Another efficiency argument is that the manufacturer may use an exclusive dealing arrangement in order to maintain the value of his product. It may want to exclude certain "inferior" retailers from selling its product if these retailers are not sufficiently trained or if their reputation does not fit the reputation of the product. For example, the producer of expensive luxury perfumes may not want its fragrances to be sold in cheap discount stores because this damages its reputation.

An exclusive dealing contract may also be used to prevent "excessive" entry. If entry involves significant sunk costs, there may be too much entry because of the "business stealing effect": A new firm that contemplates entering the market does not take into account that some of its prospective customers will switch away from existing firms. Thus the revenue that the new firm generates may be larger than the social value it creates. Other efficiency defences are parallel to those given for a refusal to deal in Section III.4.

[36] See Aghion and Bolton (1987), *op. cit.*
[37] See Rasmussen, Ramseyer and Wiley (1991) and Segal and Whinston (2000), *op. cit.*

c) Implementation: an example

We argued above that an exclusive dealing contract may have anticompetitive effects only if it imposes a negative externality on a third party. Thus, the competition authority first has to point out exactly which externality on a third party it is that explains why the buyer agreed to accept an anticompetitive exclusive dealing contract. If the competition authority cannot identify such an external effect, then there is no anticompetitive effect and the case should be dismissed.

An example of such an externality is given by what are known as "buyers' coordination stories".[38] Consider a potential entrant which needs a minimum amount of demand for entry to be profitable (e.g. because of large fixed costs). Suppose that there are many buyers and that the incumbent seller offers to each of the buyers a (small) rebate if he signs an exclusive dealing contract. A buyer accepts this offer if the rebate is larger than his expected loss due to the reduced likelihood of entry. However, he does not take into account the negative external effect of his decision on the profits of other buyers. Thus, if buyers cannot coordinate their behaviour, the externality they impose on each other may result in deterring entry. Exclusive dealing contracts can thus be anticompetitive when buyer coordination is needed.

In order to establish that such a mechanism is at work, the competition authority should prove a number of facts:

- there are many small size buyers;
- no communication nor coordination between them on procurement of input is at work. For instance, this would not be verified where super-markets delegate the power of listing products to a centralized agency and even negotiate the transaction on aggregate sales with this agency;
- on the supply side, the magnitude of the efficient size for an entrant should also be evaluated. If it appears that a competitor can enter the market selling a small number

[38] See, for example, Aghion and Bolton (1987), *op. cit*., Rasmussen, E. B., J. M. Ramseyer and J. S. Wiley (1991), Naked Exclusion, *American Economic Review*, 81: 1137-1145, Segal, I., and M. D. Whinston (2000), Naked Exclusion: Comment, *American Economic Review*, 90(1):296-309, and Fumagalli, C., and M. Motta (2002), Buyers' coordination exclusive dealing and entry, when buyers compete, London CEPR Discussion paper.

of units, for instance because the required fixed costs are small, then exclusive dealing contracts are unlikely to prevent entry. On the other hand, exclusive dealing contracts are more likely to be anticompetitive if large fixed costs are at work.

The next step for the competition authority is to check whether there are efficiency defences justifying the exclusive dealing arrangement? Even if there is a negative external effect on a third party, it is still possible that the exclusive dealing contract gives rise to efficiencies that outweigh the anticompetitive effects. The competition authority should intervene only if this possibility can be ruled out.

Section 6: Predatory Pricing

a) Potential anticompetitive effects

Predatory pricing is a strategy that consists of two phases: a predatory phase and a recoupment phase. During the predatory phase a dominant firm engages in a price reduction, which is profitable only because it eliminates or sufficiently weakens the competitive conduct of the firm's rival(s) or potential rival(s) during the recoupment phase. Thus, predatory behavior is characterized by a phase of short-term price reduction succeeded by a phase during which the predator, equipped with sufficiently strong market power, can raise prices so as to increase long-run profits once competition is eliminated or sufficiently weakened. As mentioned in Chapter 2, we can distinguish at least three types of strategic models of predatory pricing: (1) predation associated with imperfections in the financial markets; (2) signaling models; and (3) reputation models.

b) Pro-competitive effects and efficiency considerations

In general, predatory pricing poses a challenge of intrinsic difficulty to the antitrust authorities, because the challenged practice is that of a low price during the predatory phase. But at a general level the objective of antitrust is precisely to promote competitive prices. Furthermore, the Chicago critique seriously questioned whether predatory pricing could emerge as an equilibrium phenomenon consistent with rational behavior on behalf of firms. For example, the deep pocket theory underlying financial predation was challenged by arguments according to which a perfectly competitive capital market would finance a financially constrained prey, and thereby eliminate the future returns from the incumbent's investments into predation. However, as subsequent developments in industrial organization have established, capital market imperfections would constitute circumstances for perfectly rational financial predation.

Overall, the challenge for the design of antitrust policy against predation is related to the ability of the antitrust authority to separate a price that is low for predatory purposes from a price that might be set very low as part of an efficiency-enhancing competitive process. For example, if there are switching costs, network effects, product complementarities or learning effects it could be perfectly legitimate and consistent with healthy competition that firms set very low prices when they are introducing new products, when they are targeting new customer segments or rivals' installed bases, or when they are in the first phase of the learning curve. Thus, the competition authority with limited knowledge of industry- and firm-specific data faces a complex problem when attempting to identify those circumstances under which loss-inducing predatory prices cause harm to competition. For that reason the antitrust authorities have to be fully aware of the risks of misclassification when approaching a predation case.

c) Implementation: An example

We now apply the effects-based approach to predatory pricing cases. As for the other practices, the first priority should be to identify the nature of the exclusionary effect that is considered. Chapter II discussed several alternative scenarios; for the sake of exposition, we will focus here on test market predation based on signal jamming.

Consider a firm that reflects on whether to enter a given market, and assume that there is substantial uncertainty regarding the profitability of entry. Rather than to commit itself to a highly risky entry decision this firm, the prey, acquires information by entering a test market, i.e. a limited product or geographic market, with the idea of learning the demand so as to be able to assess whether full-scale entry is profitable or not. The incumbent firm, the predator, responds to this entry, or threat thereof, by lowering its price in the test market in order to prevent the prey from learning about demand under normal competitive conditions. The predator's price in the test market, however, differs from its price conduct in other markets where the predator only faces established competition.

In contrast to financial predation, test market predation based on signal jamming does not necessarily require any informational asymmetry between the prey and outside investors, or any other source of credit market imperfection. The crucial feature is that the predator jams market signals in such a way that the prey is unable to form a reliable estimate of demand. Such signal jamming may lead to complete entry deterrence or delayed entry and thus harm competition in that way. Proof of test market predation therefore essentially requires evidence that the predatory price reduction prevents the prey from learning about demand under normal competitive conditions, without restricting the instruments available to the entrant for information acquisition.[39]

In particular, the recoupment ability of the predator has to be carefully assessed. The antitrust authority must explain why the prey's exit or deterred entry causes persistent competitive harm. In order to convincingly establish test market predation based on signal jamming, alternative channels of acquiring information must be unavailable, significantly more expensive or less accurate: if competitive alternative channels of information acquisition are unavailable, predatory test market pricing would constitute a strategic entry barrier that could exclude rivals from the market in a persistent way and thereby harm competition; if instead signal jamming strategies fail to exclude other firms under similar market conditions the predator would not be able to recoup and the conditions of persistent competitive harm would not be satisfied.

Overall sensible policy recommendations regarding predation should require that the antitrust authority is able to prove by reference to convincing evidence that the critical assumptions of a well-defined strategic model of predation are satisfied and that the conduct of the predator is consistent with that theory. In this report we have characterized sets of critical assumptions

[39] Bolton, Broadley and Riordan (2000) present a detailed application of the effects-based approach to such test market predation in a case involving entry into the coffee market in Eastern USA in the 1970's. In this market General Foods engaged in substantial price cuts in particular test market areas, which the entrant, Procter & Gamble, had selected in order to elucidate the demand for the coffee brand (Folgers) it wanted to introduce in Eastern USA.

for predation to potentially offer a consistent and credible story of competitive harm in the cases of predation based on signal jamming (this section) or based on capital market imperfections (Section II.1). In both cases, identifying the nature of the concern (a consistent "predation story") allows the antitrust authority to spell out the key facts that need to be established, such as the need for the prey to acquire market information or to attract outside investors.[40] It can be noted that these key facts had little to do with the predator's cost; rather, they concerned the target's ability to get access to finance or reliable information. Thus, the price-cost rules currently used in EU case law do not necessarily provide convincing evidence that the critical facts of a sound predation story are present. Finally, it is up to the antitrust authority to make certain that these key facts are indeed present. If it has done so and if the evidence of predatory pricing is sufficiently strong, the burden of proof for a convincing efficiency defense of its conduct should switch to the defendant.

[40] As these two forms of predation make clear, the nature of the facts that need to be established indeed critically depends on the predation story involved.